

PERSPECTIVES ON INDIVIDUAL DIFFERENCES

Theoretical Foundations of Behavior Therapy

Edited by
Hans J. Eysenck
and Irene Martin

Theoretical Foundations of Behavior Therapy

PERSPECTIVES ON INDIVIDUAL DIFFERENCES

CECIL R. REYNOLDS, *Texas A&M University, College Station*
ROBERT T. BROWN, *University of North Carolina, Wilmington*

DETERMINANTS OF SUBSTANCE ABUSE
Biological, Psychological, and Environmental Factors
Edited by Mark Galizio and Stephen A. Maisto

**HISTORICAL FOUNDATIONS OF EDUCATIONAL
PSYCHOLOGY**
Edited by John A. Glover and Royce R. Ronning

THE INDIVIDUAL SUBJECT AND SCIENTIFIC PSYCHOLOGY
Edited by Jaan Valsiner

THE NEUROPSYCHOLOGY OF INDIVIDUAL DIFFERENCES
A Developmental Perspective
Edited by Lawrence C. Hartlage and Cathy F. Telzrow

PERSONALITY AND INDIVIDUAL DIFFERENCES
A Natural Science Approach
Hans J. Eysenck and Michael W. Eysenck

PERSONALITY DIMENSIONS AND AROUSAL
Edited by Jan Strelau and Hans J. Eysenck

PERSPECTIVES ON BIAS IN MENTAL TESTING
Edited by Cecil R. Reynolds and Robert T. Brown

THEORETICAL FOUNDATIONS OF BEHAVIOR THERAPY
Edited by Hans J. Eysenck and Irene Martin

A Continuation Order Plan is available for this series. A continuation order will bring delivery of each new volume immediately upon publication. Volumes are billed only upon actual shipment. For further information please contact the publisher.

Theoretical Foundations of Behavior Therapy

Edited by
Hans J. Eysenck
and
Irene Martin

Institute of Psychiatry
London, England

Springer Science+Business Media, LLC

Library of Congress Cataloging in Publication Data

Theoretical foundations of behavior therapy.

(Perspectives on individual differences)

Includes bibliographies and index.

1. Behavior therapy. I. Eysenck, H. J. (Hans Jurgen), 1916- . II. Martin, Irene.

III. Series. [DNLM: 1. Behavior Therapy. WM 425 T3958]

RC489.B4T516 1987

616.89/142

87-17168

ISBN 978-1-4899-0829-2

ISBN 978-1-4899-0827-8 (eBook)

DOI 10.1007/978-1-4899-0827-8

© 1987 Springer Science+Business Media New York

Originally published by Plenum Press, New York in 1987.

Softcover reprint of the hardcover 1st edition 1987

All rights reserved

No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise, without written permission from the Publisher

Contributors

Chris R. Brewin MRC Social Psychiatry Unit, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, England

Anthony Dickinson Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, England

Georg H. Eifert Department of Behavioural Sciences, James Cook University of North Queensland, Townsville, Queensland 4811, Australia

Ian M. Evans Department of Psychology, State University of New York at Binghamton, Binghamton, New York 13901

H. J. Eysenck Institute of Psychiatry, De Crespigny Park, London SE5 8AF, England

Michael W. Eysenck Department of Psychology, Birkbeck College, University of London, London WC1 E 7HX, England

J. A. Gray Department of Psychology, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, England

Leslie S. Greenberg Department of Psychology, York University, North York, Ontario, M3J 1P3, Canada

R. S. Hallam Audiology Centre, Institute of Laryngology and Otology, Gray's Inn Road, London WC1X 8DA, England

P. J. Higson Department of Psychology, North Wales Hospital, Denbigh, Clwyd LL16 5SS, Wales

P. J. Horne Department of Psychology, University College of North Wales, Bangor, Gwynned LL57 2DG, Wales

Michael J. Kelley University of Hawaii, 1960 East-West Road, Honolulu, Hawaii 96822

A. B. Levey MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, England

Brett T. Litz Department of Psychology, State University of New York at Binghamton, Binghamton, New York 13901

C. F. Lowe Department of Psychology, University College of North Wales, Bangor, Gwynned LL57 2DG, Wales

Irene Martin Department of Psychology, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, England

Andrew Mathews Department of Psychology, St. George's Hospital Medical School, Cranmer Terrace, London SW17, England

Susan Mineka Department of Psychology, University of Texas, Austin, Texas 78712

Kieron P. O'Connor Hôpital Louis H-Lafontaine, Université de Montréal, Montréal P.Q. H1N 3MS, Canada

M. J. Power MRC Social Psychiatry Unit, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, England

Jeremy Safran Clarke Institute of Psychiatry, Toronto, Ontario M5T 1RH, Canada

Angela Schorr Universität Eichstätt, Ostenstr. 26-28, 8078 Eichstätt, West Germany

J. M. G. Williams MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, England

Preface

In this book we have attempted to confront a number of issues that are intimately related to the theoretical basis of behavior therapy. We believe that behavior therapy is an extremely efficient procedure for the treatment of neurotic disorders; that it is based on certain principles derived from learning theory; and that it is unique in using basic scientific principles in psychology in the service of applied and practical ends. We believe that we are here dealing with much more than the advantageous use of serendipitous borrowings from nonexistent principles, the cookbook collection of precepts, methods, and working rules that happen to have lasting effects. We also believe that there is truly a general principle underlying behavior therapy, rather than a varied mass of nonintegrated therapies that have little in common other than a name.

These beliefs are often contested, but usually those who oppose them do so on the basis of misconceptions and misunderstandings that indicate a lack of knowledge of fundamental facts. It is the purpose of this book to remove these misconceptions and misunderstandings, and to bring up to date our knowledge in certain fundamental areas of learning theory, behavior therapy, and the biological foundations of personality and individual differences.

There are three major groups of misconceptions and misunderstandings. The first of these relates to beliefs held by many psychiatrists and cognitive psychologists relating to behavior therapy. As Wolpe (1986) has pointed out, although a task force of the American Psychiatric Association (1973) reached the conclusion that behavior therapy has "much to offer in the service of modern clinical and social psychiatry," it is not widely taught or used by the profession, and most psychiatrists believe that its applicability is limited to certain circumscribed problems, such as phobias, and that it is simplistic and irrelevant to the complexities of the total personality. Wolpe goes on to note that misreporting, often with pejorative overtones, has been the rule ever since (Locke, 1971; Peterson, 1966; Rotter, 1959). Others, like Frank (1973), Bergin (1971), Strupp (1978), and Garfield (1981) "have played down behavior therapy and have consistently ignored positive data" (Wolpe, 1986, p. 191).

There is a widespread belief that behavior therapists "assume that what goes on subjectively within the patient is irrelevant and that all that matters is how he behaves" (Marmor & Woods, 1980). Similarly, Goisman (1985) looks upon behavior techniques as adjuvant exercises whose effects depend on reward and punishment! Cognitive behavior therapists, like Beck (1976), Ellis (1974), and Mahoney (1977),

declare that behavior therapy is simple and mechanistic, and that standard behavior therapy overlooks thoughts and feelings. Thus Beck (1976) has stated explicitly that behavior therapists “selectively exclude information regarding patients’ attitudes, beliefs and thoughts.” As Wolpe (1986) has pointed out in some detail, and as we shall see in the course of this book, these notions are erroneous and have no substance in fact.

They do, however, mirror a second set of misconceptions and misunderstandings held by many therapists concerning learning theory. Thus it is assumed that learning theory, particularly in the form given to it by Skinner, disregards “covert events,” largely on the ground that there can be no public agreement about their validity. This is indeed the view originally held by methodological behaviorism, but Skinner has consistently argued against this view on the grounds that it misguidedly adheres to the outmoded tenets of logical positivism and operationism; nowadays it would be very difficult to find learning theorists who advocate such a view. As will be shown later in the book, the principal distinguishing feature of Skinner’s radical behaviorism is that he considers that a science of behavior, like other sciences, must deal with events that are not directly observable, and that inference consequently is essential in the study of behavior (Skinner, 1945, 1974). As Lowe, Horne, and Higson point out in Chapter 7 of this book.

It is surely a strong irony of contemporary psychology that an approach which, as far back as 1945, established its identity on the basis of its recognition of the inner life of humans should so often be charged with the error of ruling it out of court.

In a similar way, Pavlovian conditioning and the laws that have been found to govern it are often held to be “mechanistic” and lacking in cognitive content. This may be true of Watson’s very primitive form of behaviorism, but it could certainly not be said to be true of Pavlov, with his “second signaling system.” Such criticism is simply a parody of modern theories that have abandoned the S-R model and are almost exclusively of the S-S type (Zuriff, 1985). It is a curious feature of psychology that we should still be fighting battles against opponents long since slain, and that we refuse to deal with the much more adequate theories advanced in more recent years!

Last but not least, learning theorists, behavior therapists, psychiatrists, and cognitive psychologists underestimate the importance of genetics and the relevance of individual differences for the origin and treatment of neurotic disorders, and have ideas about behavioral genetics that are well behind the times. Here too, therefore, we have the odd situation that battles are being fought over issues—and in terms—that are 50 years out of date. It is curious that the development of a unified science of psychology (Eysenck, 1985) should be so needlessly delayed by the refusal of many leading psychologists to acquaint themselves with the latest position in neighboring and relevant fields to those in which they themselves are expert. However that may be, our hope is that this book will bring together exponents of the various disciplines mentioned and will lead to a greater understanding on each side of what the other side is saying, what are in fact the major theoretical preconceptions of the age, and how best to integrate the many strands that make up this very diverse and intriguing field.

REFERENCES

- American Psychiatric Association. (1973). *Task force report: Behavior therapy in psychiatry*.
- Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.
- Bergin, A. E. (1971). The evaluation of therapeutic outcome. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis*, (pp. 217–270). New York: Wiley.
- Ellis, A. (1974). *Humanistic psychotherapy: The rational-emotive approach*. New York: Julian.
- Eysenck, H. J. (1985). The place of theory in a world of facts. *Annals of Theoretical Psychology*, 3, 17–72, 103–114.
- Frank, J. D. (1973). *Persuasion and healing* (Rev. ed.). Baltimore, MD: John Hopkins University Press.
- Garfield, S. L. (1981). Psychotherapy: A 40-year appraisal. *American Psychology*, 36, 174–183.
- Goisman, R. M. (1985). The psychodynamics of prescribing in behavior therapy. *American Journal of Psychiatry*, 142, 675–679.
- Locke, E. A. (1971). Is behavior therapy behavioristic? *Psychological Bulletin*, 76, 318.
- Mahoney, M. J. (1977). Reflections on the cognitive-learning trend in psychotherapy. *American Psychologist*, 32, 5–13.
- Marmor, J., & Woods, S. M. (1980). *The interface between psychodynamic and behavior therapy*. New York: Plenum Press.
- Patterson, C. H. (1966). *Theories of counselling and psychotherapy*. New York: Harper & Row, 1966.
- Rotter, J. B. (1959). Substituting good behavior for bad. *Contemporary Psychology*, 4, 176–178.
- Skinner, B. F. (1945). The operational analysis of psychological terms. *Psychological Review*, 52, 270–277.
- Skinner, B. F. (1974). *About behaviorism*. London: Jonathan-Cape.
- Strupp, H. (1978). Psychotherapy research and practice: An overview. In M. Lambert & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change*, (pp. 218–254). New York: Wiley.
- Wolpe, J. (1986). Misrepresentation and under employment of behavior therapy. *Comprehensive Psychiatry*, 27, 192–200.
- Zuriff, G. E. (1985). *Behaviorism: A conceptual reconstruction*. New York: Columbia University Press.

Contents

PART I INTRODUCTION

Chapter 1

Behavior Therapy 3

H. J. Eysenck

Behavior Therapy and Psychotherapy	3
The Concept of Neurosis	4
Criticisms of Behavior Therapy	9
Claims of Cognitive Behavior Therapists	12
A New Conditioning Model of Neurosis	16
The Neurobiology of Incubation of Fear/Anxiety	19
Is Exposure a Necessary Condition for Fear Reduction?	23
Summary and Conclusions	28
References	29

Chapter 2

Behavior Modification 37

Angela Schorr

Definition of Terms and Fundamental Principles	37
Practical Application: A One-Way Street? Science and Technology in Applied Behavior Analysis	40
The Token Economy: An Achievement-Oriented Society Made in the Behaviorist Laboratory	45
Conclusion	50
References	52

PART II CONDITIONING THEORY

Chapter 3

Animal Conditioning and Learning Theory 57

Anthony Dickinson

Cognition and Conditioning	58
Reinforcer Reevaluation	58
Mediated Conditioning	60
Behavioral Autonomy	61
Conditions of Acquisition	62
Event Correlation	62
Selective Conditioning and Blocking	64
Learned Irrelevance	66
Causal Relevance	66
Long-Delay Conditioning	67
Theories of Conditioning	68
The Wagner Theory	68
The Pearce-Hall Theory	69
The Mackintosh Theory	71
Inhibitory Conditioning	72
Occasion Setting	73
Instrumental Conditioning	75
Conclusions	77
References	78

Chapter 4

A Primate Model of Phobic Fears 81

Susan Mineka

Introduction	81
Symptomatology of Primate Fear of Snakes	83
Parallels with Human Phobias Using Lang's Three-Systems	
Model of Fear	83
Therapy for Fear of Snakes	85
Parallels in Treatment Outcome following Flooding	
Therapy	85
Etiology through Observational Conditioning	88
Immunization/Prevention of Fears	103
Overview and Conclusions	106
References	109

Chapter 5

Evaluative Conditioning: A Case for Hedonic Transfer 113

A. B. Levey and Irene Martin

Classical Conditioning of Preferences	114
An Early Prototype	115
The Evaluative Conditioning Paradigm	115
The Theory of Evaluative Conditioning	117
Classical Conditioning Concepts	117
Antecedents of the Evaluative Paradigm	118
A Theoretical Framework	121
Clinical Applications	124
Relevance to Therapy	124
Treatment Techniques	126
Clinical Assessment	127
References	129

Chapter 6

Knowledge, Action, and Control 133

Irene Martin and A. B. Levey

Knowledge, Action, and Control	133
Statement of the Problem	133
Cognition and Conditioning	134
Verbalizable Knowledge and Awareness	135
Effects of Verbalizable Knowledge on Conditioning Performance	136
Autonomic Conditioning and Stimulus Awareness	137
The Eyelid Conditioning Paradigm	138
Effects of Verbalizable Knowledge on Cognitive Performance	143
Information Load and Salience	144
Alternative Theoretical Models	145
Integration of Behavior Stimulus Control and Mental Representations	146
The Architecture of Cognition	147
Clinical Conditions	147
Conclusions	149
References	150

Chapter 7

Operant Conditioning: The Hiatus between Theory and Practice in Clinical Psychology 153

C. F. Lowe, P. J. Horne, and P. J. Higson

Introduction	153
Behavioral Theory in Practice	156

Token Economies	156
Contingency Management	158
Conclusion	160
References	161

Chapter 8

Language Conditioning: Clinical Issues and Applications in Behavior Therapy 167

Georg H. Eifert

Introduction	167
Terminology and Basic Concepts	168
Language Conditioning in the Social-Behaviorist Paradigm	169
Language Conditioning and the Functions of Language	169
Language Conditioning in Anxiety and Depression	170
Applications of Language Conditioning in Behavior Therapy	171
Semantic and Verbal Conditioning Interventions	172
Language Conditioning in Cognitive-Behavioral Interventions	174
Effects of Language Conditioning on Verbal, Affective, and Behavioral Changes	178
Modifying Clients' Verbal Behavior in Therapy: An Analogue to Psychotherapy?	178
Language Conditioning and Emotional Arousal	179
Language Conditioning and Behavioral Changes	180
Conclusions	181
Variables Influencing the Effectiveness of Language Conditioning	183
Personality Characteristics and Interpersonal Skills of the Therapist	183
Personality Characteristics of Clients	184
Client Knowledge of Conditioning Contingencies and Countercontrol	187
Conclusions	188
References	190

PART III COGNITIVE THEORY

Chapter 9

Trait Anxiety and Cognition 197

Michael W. Eysenck and Andrew Mathews

Introduction	197
The Cognitive System: Overview	199
The Cognitive System: Performance Deficits	200
The Cognitive System: Long-Term Memory	203
The Cognitive System: Reactions to Threat	206

Toward a Cognitive Model	213
References	213

Chapter 10

Clinical Anxiety and Cognition 217

Andrew Mathews and Michael W. Eysenck

Introduction	217
The Origins of Phobias and Anxiety States	218
A Cognitive Model of Clinical Anxiety	219
Cognitive Content in Anxiety Disorder	221
Attention and Clinical Anxiety	222
Recall and Recognition in Clinical Anxiety	225
Judgment and Interpretation in Clinical Anxiety	226
Clinical Implications of Cognitive Research	228
References	232

Chapter 11

Cognitive Theories of Depression 235

M. J. Power

Introduction	235
Three Theories	236
Reformulated Learned Helplessness	236
Semantic Network Theories	240
Beck's Cognitive Therapy	243
Some Speculations	245
A Putative Cognitive Basis	245
Vulnerability to Depression	249
The Process of Depression	251
Conclusions	252
References	253

Chapter 12

Cognitive Treatment of Depression 257

J. M. G. Williams

Introduction	257
Analysis of Subcomponents of Treatment Package	261
Predictors of Treatment Response	265
Process Research	266
Cognitive Psychology and Depression	268

Concluding Remarks	271
References	273

Chapter 13

Cognitive Theories of Motivation 277

Chris R. Brewin

Introduction	277
Two Cognitive Systems	279
Expectancies	283
Values and Goals	286
Clinical Implications	289
References	291

Chapter 14

Emotion, Cognition, and Action 295

Leslie S. Greenberg and Jeremy Safran

Introduction	295
Associative Network Models	296
Schema Models	299
The Evaluative Conditioning Model	299
Comparison of Models	301
An Integrative Model of Emotional Processing	301
The Emotional Synthesis Process	301
Adaptive Action Tendencies	303
Specific Patterning and Somatic Feedback	304
Higher Level Processing	305
Emotion in Psychotherapy and Behavior Change	306
Clinical Implications	306
Integrated Intervention	307
Emotional Change Processes	309
Conclusion	310
References	310

PART IV

DESCRIPTION AND THE ORGANIZATION OF BEHAVIOR

Chapter 15

Prospects for Theoretical Progress in Behavior Therapy 315

R. S. Hallam

Introduction	315
The Changing Role of Theory in Behavior Therapy	316

Technical and Theoretical Eclecticism 316
 Behavior Therapy as Applied Science 318
 What is Behavioral about Behavior Therapy? 321
 Mentalism and Cognitivism: A Behavioral Response 322
 Behaviorism and the Content of Behavior 326
 Conclusion 327
 References 327

Chapter 16

Behavioral Assessment: A New Theoretical Foundation for Clinical Measurement and Evaluation 331

Ian M. Evans and Brett T. Litz

Introduction 331
 Behavioral Assessment and Measurement Theory 333
 Reliability 333
 Validity 334
 Improving the Veridicality of Measurement 336
 Cognitive Personality Assessment 339
 Personality and the Purpose of Assessment 341
 Target Behavior Selection 342
 Clinical Decision Making 344
 Integration and Conclusion 347
 References 349

Chapter 17

A Response Process Model of Behavior 353

Kieron P. O'Connor

Introduction 353
 The Problem of Behavior 354
 Commonsense Definitions of Everyday Behavior 354
 Behavior as Representation 355
 Sensory Models of Behavior 357
 Conditioning Models of Behavior 359
 Constitutional Approaches to Behavior 361
 Physiological Theories 361
 Cognitive Theories 362
 Behavior as Response Controlled 363
 Behavior as Self-Regulated Action 364
 Behavior as a Relational Act 364
 Behavior as an Ecological Act 365
 The Response-Centered Approach to Behavior 367
 Response Process Analysis and Conventional Functional Analysis 369
 Clinical Example of Process Analysis 371

Concluding Remarks	373
References	374

PART V BIOLOGICAL BASES OF PERSONALITY AND BEHAVIOR

Chapter 18

The Role of Heredity, Environment, and “Preparedness” in the Genesis of Neurosis 379

H. J. Eysenck

Conclusions	394
References	398

Chapter 19

Hormones and Clinical Anxiety: An Imbalanced Neuromodulation of Attention 403

Michael J. Kelley

Introduction	403
Some Basic Concepts	404
Definitions of Conditioning, Attention, and Neuroses	405
A Definition of Neurohormones	406
Neurohormones and the Extinction of Associations in Animals	407
Corticotropin and Related Peptides	407
Cortisol and Corticosterone	409
Corticotropin-Releasing Factor	410
Beta-Endorphin	410
Vasopressin and Oxytocin	411
Neurohormones and the Incubation of Anxiety in Animals	411
Neurohormones, Attention, and Anxiety	414
Neurohormones and Attention in Animals	414
Neurohormones and Attention in Humans	415
Attention and Human Anxiety	417
Neurohormones and Human Neuroses	417
The Interpretation of Samples of Stress Hormones	418
The Role of Vasopressin in Neuroses	419
The Role of Opioid Receptors in Incubation	420
Neurohormones and Therapy for Neuroses	423
Neurohormones and the Treatment of Neuroses	423
Behavior Therapy and Neurohormones	423
Summary: Toward a Clinical Neuroscience of Anxiety	425
References	425

Chapter 20

Interactions between Drugs and Behavior Therapy 433

J. A. Gray

The Partial Reinforcement and Partial Punishment Effects	433
The Effects of Antianxiety Drugs on the PREE and PPE	435
Clinical Implications	439
A More Positive Note	441
References	446

**PART VI
CONCLUSION**

Chapter 21

**Concluding Comments on Theoretical Foundations and Requirements
in Behavior Therapy 451**

Irene Martin

Introduction	451
The Conditioning View of Behavior	453
Cognitive Models	456
Behavior Therapy: Models and Views of Behavior	460
References	464

Index 465

PART I

INTRODUCTION

CHAPTER 1

Behavior Therapy

H. J. Eysenck

BEHAVIOR THERAPY AND PSYCHOTHERAPY

The term *behavior therapy* has been in use for no more than 30 years, having been introduced to mark a Kuhnian revolution (Barnes, 1982; Kuhn, 1959, 1970, 1974) in the prevailing theories of neurosis (Eysenck, 1959, 1960, 1964). At the time, the prevailing paradigm was a Freudian, psychodynamic one, and it may be argued that since then there has been a paradigm shift of fundamental importance to psychiatry and clinical psychology (Eysenck, 1985, 1987). This paradigm shift from Freud to Pavlov, from psychotherapy to behavior therapy, from emotional insight learning to Pavlovian extinction and deconditioning, is in large part based on the recognition that Freudian theory has essentially failed to produce methods of treatment superior to placebo treatment, or even to no treatment at all (Eysenck, 1952; Hattie, Sharpley, & Rogers, 1984; Prioleau, Mardock, & Brody, 1983; Rachman & Wilson, 1980). Meta-analysis (Shapiro & Shapiro, 1982; Smith, Glass, & Miller, 1980) has been suggested to provide evidence in favor of the effectiveness of psychotherapy, but the method itself has been severely criticized (Eysenck, 1983; Matt & Wittman, 1985; Searles, 1985) and in any case Smith, Glass, and Miller completely failed to show that any of the theories examined (with the exception of behavior therapy) had any *specific* effects, that is, effects traceable to the special theory on which the therapy was based. Furthermore, they failed to compare psychotherapy with placebo treatment, but used placebo treatment instead as one of the 18 treatments examined! Even worse, from the point of view of psychotherapy and psychoanalysis, is the fact that strong negative effects of these types of treatment have been found (Hadley & Strupp, 1976; Strupp, Hadley, & Gomes-Schwartz, 1977).

Eysenck (1959) suggested 10 major points on which the new paradigm differs from the old; these are given in Table 1. In spite of many criticisms, it can still be maintained that these points encapsulate the new paradigm, and that work done over the past 30 years has shown that, in the words of Lakatos (1970), it constitutes a

TABLE 1

Psychotherapy	Behavior therapy
1. Based on inconsistent theory never properly formulated in postulate form.	Based on consistent, properly formulated theory leading to testable deductions.
2. Derived from clinical observations made without necessary control observations or experiments.	Derived from experimental studies specifically designed to test basic theory and deductions made therefrom.
3. Considers symptoms the visible upshot of unconscious causes ("complexes").	Considers symptoms as unadaptive conditioned responses.
4. Regards symptoms as evidence of <i>repression</i> .	Regards symptoms as evidence of faulty learning.
5. Believes that symptomatology is determined by defence mechanisms.	Believes that symptomatology is determined by individual differences in conditionability and autonomic lability, as well as accidental environmental circumstances.
6. All treatment of neurotic disorders must be <i>historically</i> based.	All treatment of neurotic disorders is concerned with habits existing at <i>present</i> ; their historical development is largely irrelevant.
7. Cures are achieved by handling the underlying (unconscious) dynamics, not by treating the symptom itself.	Cures are achieved by treating the symptom itself, i.e., by extinguishing unadaptive C.Rs and establishing desirable C.Rs.
8. Interpretation of symptoms, dreams, acts, etc. is an important element of treatment.	Interpretation, even if not completely subjective and erroneous, is irrelevant.
9. Symptomatic treatment leads to the elaboration of new symptoms.	Symptomatic treatment leads to permanent recovery provided autonomic as well as skeletal surplus C.Rs are extinguished.
10. Transference relations are essential for cures of neurotic disorders.	Personal relations are not essential for cures of neurotic disorder, although they may be useful in certain circumstances.

progressive research program or problem shift, whereas the psychoanalytic theory has proved to be a degenerative research program. In this opening chapter, it may be appropriate to discuss some of the criticisms that have been made, some of the advances that have been recorded, and some of the changes in theory that have been suggested.

In recent years there has been an attempt to bring together these two incompatible models of neurotic behavior and treatment (e.g., Goldfried, 1980; Wachtel, 1977), but these attempts have not found favor in the eyes of critics like Franks (1984), Messer and Winokur (1980), and Yates (1983), who saw more clearly the incompatibility of the two approaches. Wolpe (1981) also clearly saw the irreconcilable differences between behavior therapy and psychoanalysis, and argued against the pseudoscientific eclecticism that would result from any attempt to merge them.

THE CONCEPT OF NEUROSIS

We may start with the observation that behavior therapy is intimately related with the concept of neurosis. It attempts to explain the occurrence of neurotic disorders, and it attempts to suggest methods of treatment of neurotic disorders. It thus

inevitably confronts Mowrer's (1948, 1950) paradox, which he identified by pointing, as a central feature of neurosis, to the fact that the self-defeating behavior of the neurotic is self-perpetuating. As he put it: "the neurotic paradox lies in the fact that human behavior is sometimes indefinitely perpetuated despite the fact that it is seriously self-defeating" (Mowrer, 1950, p. 524). The explanation of the Mowrer paradox given by Watson and Rayner (1920) is of course in terms of Pavlovian conditioning; neurotic symptoms are conditioned emotional and skeletal autonomic and behavioral responses that are immune from rational criticism. As I have pointed out elsewhere (Eysenck, 1968), conditioning theory so conceived does not explain why there is no extinction when self-defeating behaviors are found to be self-punishing, and are not reinforced (Kimmel, 1975). Watson's theory, although along the right lines, requires careful restatement in the light of more recent experiments and theoretical developments.

Attempts have been made to find an operant explanation of Mowrer's neurotic paradox (Tryon, 1978). This theory is based on Woods's (1974) taxonomy of instrumental conditioning. His analysis attempts to explain the apparent contradiction of the law of effect by postulating a particular subset of eight binary combinations of response contingencies, where one contingency accelerates response omission whereas the other decelerates response omission. The theory is implausible because of its complexity, because it has no factual support, but mainly because it fails to account for many well-established features of the development of neurotic responses, such as their insidious onset and incrementation through CS-only exposure (Eysenck, 1979).

Mowrer's paradox does not define neurosis, and in recent years there has been an attempt by psychiatrists to get rid of the term altogether, as for example in DSM-III. This official manual of psychiatric diagnosis has been extensively reviewed by Eysenck, Wakefield, and Friedman (1983), who point out that DSM-III is based on no particular empirical evidence, but is merely an attempt to resolve differences between psychiatric schools by committee decisions. What has happened is simply a substitution of many neurotic disorders, often highly correlated, for a general term, neurosis, under which these different neurotic disorders could be subsumed (Gossop, 1981). To say this is not to deny that the term may be difficult to define, or that it is useful to subdivide neurotic disorders into subsets that can be independently classified and diagnosed. Nevertheless, as we shall show, it appears to be true that there is a large number of mental disorders characterized by anxiety and other similar mental/autonomic/behavioral responses to stimuli that are normally unlikely to lead to strong and lasting responses of this type. These strong and lasting emotional responses in turn lead to behaviors (such as obsessive-compulsive handwashing) the purpose of which is to reduce neurotic anxiety (Gossop, 1981). Watson's theory posits that these emotional reactions are produced by Pavlovian conditioning, and may be cured by Pavlovian extinction.

Neurotics have certainly been with us for a very long time (Simms, 1985), and they impose a great stress on society (Simms, 1983). The difficulty introduced by varying levels of neurotic disorder, many of which are never seen by the psychiatrist, has given rise to a very useful model for mental illness in the community, introduced by Goldberg and Huxley (1980). They use the concepts of levels and filters. A large number of people in the community suffer from psychological symptoms in any one year (Level 1). Most of these pass the first filter and seek help from their general

practitioners (Level 2). Many of these the general practitioner recognizes as suffering from psychological symptoms (Level 3), and a quite small number of these are referred to a psychiatrist (Level 4). Of those seen as psychiatric outpatients, even fewer are admitted as inpatients (Level 5). It follows that neurotic disorders are predominantly concentrated in the community and form a smaller proportion of those patients seen by the psychiatrist. At Level 5, when the International Classification of Diseases was applied to all psychiatric inpatients in a District General Hospital Psychiatric Unit, 21% were found to be suffering from neurotic disorders (Zigmond & Simms, 1983). When serial psychiatric outpatient referrals were classified diagnostically (Level 4) 60% were found to have a primary diagnosis of neurosis (Simms & Salmons, 1975). In a large study of diagnosis in a population of 300,000 people in general practice (Level 3), the consultation rate for all neuroses was 75.5 per 1000 per annum for males, and 162.9 for females, which gave a rate of over 90% for neuroses among all psychiatric diagnoses (Office of Population Censuses and Surveys, 1974). It is clear that as we go from the community via general practice to the psychiatric outpatient/inpatient treatment, the number of neurotic patients diminishes, but the severity of their neurotic symptoms increases.

Neuroses are found extremely commonly among inpatients and outpatients of hospital specialities other than psychiatry. For instance, it has been estimated that 15% to 20% of hospital presentations in the ophthalmological clinic are for neuroses (Karseras, 1976). It seems certain that neurotic disorders constitute most of the psychiatric illnesses encountered in general practice, but that only a small proportion of these cases is referred to hospital (Kessel, 1960; Kessel & Shepherd, 1962). This large class of sufferers cannot be conjured out of existence by dropping the concept of neuroses.

The varied manifestations, symptoms, and correlates of neurotic disorder may interact with the personality dimension of extraversion-introversion to produce extraverted (hysterical) or introverted (dysthymic) disorders (Janet, 1890, 1903; Jung, 1923; Eysenck, 1947). Figure 1 shows the results of a factor analysis of various notations on 700 male neurotics, and Figure 2 shows a similar analysis of symptoms intercorrelated and factor analyzed for a large group of children in a child guidance clinic (Eysenck, 1970a). These data will illustrate the complex of feelings and behaviors constituting the different neurotic disorders. More details concerning neurosis and the personality trait of neuroticism underlying it are given in a later chapter in this book (Genetics and Preparedness), which will also discuss the relative influence of genetic and environmental determinants on neurosis.

The major evidence suggesting that it may be meaningful to postulate the concept of neurosis, as opposed to normality and psychosis, comes from factor analytic studies demonstrating (a) that these three concepts require two dimensions to accommodate the observed relationships, thus making it impossible to postulate, as Freud had done, that psychosis is merely a further development of neurotic illness, and (b) that most neurotic and psychotic disorders are more easily conceived as end points of two different continua, rather than as categorically different from normality (Eysenck, 1970b).

The theory to be outlined here states that the different types of neurotic illness arise through a process of Pavlovian conditioning, and can be eliminated through a

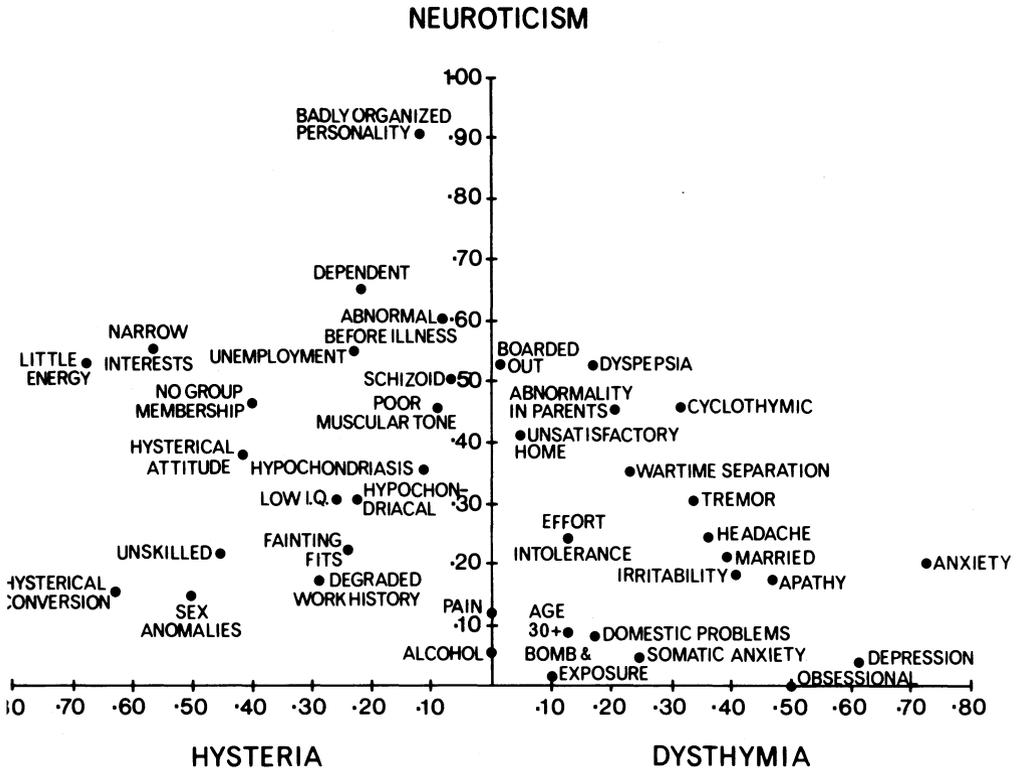


FIGURE 1. Neuroticism as a factor in neurosis, divided into introverted (Dysthymia) and extraverted (Hysteria) behaviors. (From Eysenck, 1947.)

process of Pavlovian extinction. It is often objected that there are certain types of existential problems that would be difficult to understand along these lines, and that would seem difficult to treat by means of behavior therapy. Without wishing to enter into a sustained discussion of the topic, it should be pointed out that it would be unrealistic to assume that all the patients presenting at psychiatric clinics are suffering from either neurotic or psychotic disorders, or a combination of the two. Thus, when psychotic disorders have been eliminated, it would be quite wrong to assume that the remainder must all be suffering from neurotic disorders. Not all anxieties and fears are irrational, and many children and adults presenting in psychiatric hospitals and clinics may require advice and guidance rather than behavior therapy. Similarly, if existential fears, worries, and doubts do not arise from a process of conditioning, along the lines suggested, they would not seem to fall under the general heading of neurosis. The collection of patients normally seen by psychiatrists and clinical psychologists is a relatively arbitrary grouping possessing little in the way of homogeneity. It is scientifically perfectly legitimate to subdivide this heterogeneous total into smaller groups showing that essential uniformity that is required for the elaboration of general laws.

At first sight this argument may seem circular, but in fact it is one that is commonly found in the hard sciences. If we ask whether Euclidian geometry applies

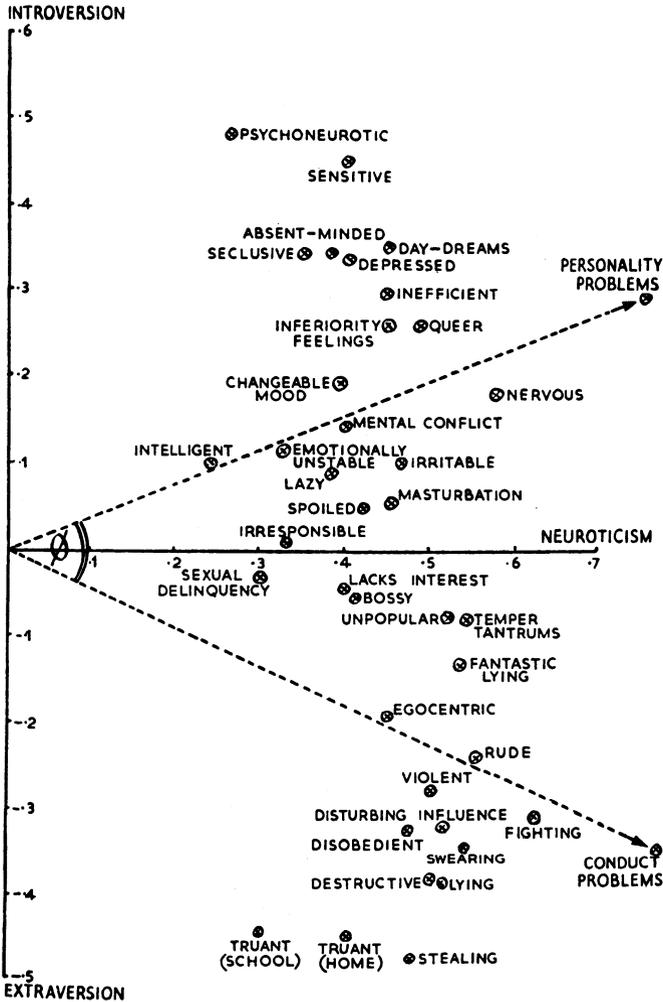


FIGURE 2. Two-factor representation of Ackerson's Correlational Study. Extraverted and introverted forms of neurotic behavior in a group of child guidance clinic children. (From Eysenck, 1970.)

to a particular type of measurement, say of a given part of the earth's surface, we answer that it applies only to planes. If we chose a small part of the earth's surface, say an acre, it is sufficiently close to a plane to make Euclidian geometry applicable. If we chose a larger surface, such as a continent, clearly this is curved, and hence Euclidian geometry does not properly apply. How do we know whether a given surface is or is not a plane? The answer of course is in terms of the application of Euclidian principles; if they apply, it is a plane, if they do not, it is not.

It is also important to remember that in our definition we are dealing with a scientific law, and that such laws always have limits to the conditions under which they apply. Take the law that tells us what the speed of fall S , in metres per second, would be at any point along the path of fall of a body dropped near the earth's surface. The formula of course is $S = 4.43^2 h$, where h is the distance that the body

has already fallen, measured in metres. Such a law would not apply to any body whose size, shape, or speed of fall is such that air resistance affects its motion appreciably. In a similar way the application of the general law relating neurosis and conditioning implies limits to the conditions under which it applies, and these should always be borne in mind. They do not constitute a negation of the law, just as little as air resistance implies a negation of the law of falling bodies.

Furthermore, the postulation of classical conditioning and extinction as being the core of the theory of neurosis does not rule out other processes (cognitive, operant conditioning, etc.) as being powerfully involved. Successful treatment along the lines of behavior therapy may produce reactions on the part of spouses, relatives, and others that reward or punish the patient for the improvement in his or her state; this may have positive or negative effects on the success of therapy. These factors, however, are adventitious rather than central, and must be sharply differentiated from the centrally placed factors postulated by the theory. More will be said on this point later on.

There has been a dearth of studies directly investigating the origins of neurotic disorders, and those that have been done have concentrated almost exclusively on phobic disorders (Murray & Foote, 1979; Öst & Hugdahl, 1981; Öhman, Dimberg, & Öst, 1985; Rimm, Janda, Lancaster, Nake, & Dittmar, 1977). The general finding seems to be that in a majority of cases conditioning experiences are remembered. In what is probably the best of these studies, Öst and Hugdahl found that vicarious experiences only accounted for 17%, and instructions/information only for 10%. Similarly Rimm *et al.* found that vicarious experiences accounted for 8% and instructions for 11%. Murray and Foote found a higher proportion of indirect ways of acquiring fears, but they did not use clinical patients with phobias, concentrating on undergraduate students with a fear of snakes. A fair number of subjects in all three studies failed to recollect any specific causal factors. "There is no clear-cut relationship between the ways of acquisition and anxiety components (subjective, behavioral, psychological), nor did the conditioning and indirectly acquired phobias differ in severity." This is clearly an area where much more research is urgently needed, preferably using interview questioning rather than questionnaire data. We shall return to it in a later section in connection with treatment procedures.

CRITICISMS OF BEHAVIOR THERAPY

We may now turn to some of the criticisms that have been made of the conception of behavior therapy here outlined. First, let us consider some points made by Breger and McGaugh (1965), Locke (1971), and London (1972). These and other authors first challenged the theoretical basis of behavior therapy, maintaining that the laws of learning on which behavior therapy was said to be based remained to be established themselves. As they pointed out, fundamental issues, such as the role of mediational events in behavior change, the nature of responses learned, and the limitations of a stimulus-response analysis, had not yet been resolved. They made the point that behavior therapy mistakenly assumed a monolithic learning theory as a basis of behavior therapy as an applied science, but, if learning theory itself had

not succeeded and resolved its major issues, then how could behavior therapy rely on the theory as an established guide (Erwin, 1978)? And in the second place, critics have argued that the principles of learning theory, if any such existed, do not in fact inform the practice of modern behavior therapists. Instead, it is argued that serendipity, nonspecific factors such as suggestibility, the personality of the therapist, etc., are responsible for the success of treatment, if any. These criticisms, which have been frequently repeated, deserve an answer.

It is certainly true that the principles of learning theory are not as firmly established as one would like (Zuriff, 1985), and that indeed the whole doctrine of behaviorism is under assault (Mackenzie, 1977). But of course this is precisely what characterizes a new paradigm. Barnes (1982) comments on the

perceived inadequacy of a paradigm as it is initially formulated and accepted . . . its crudity, its unsatisfactory predictive power, and its limited scope, which may in some cases amount to but a single application. In agreeing upon a paradigm scientists do not accept the finished product: rather, they agree to accept it as a basis for future work, and to treat as illusory or eliminable all apparent inadequacies and defects. Paradigms are refined and elaborated in normal science. And they are used in the development of further problem-solutions, thus extending the scope of scientific competences and procedures. (p. 46)

Newton's *Principia Mathematica* was dismissed out of hand by French physicists for similar reasons to those adduced by Breger and McGaugh, and his mathematical treatment of the calculus was not made rigorous until 150 years after his death, in Cauchy's *Cours d'Analyse*. If we followed the lugubrious advice of the critics, no scientific advances would ever be possible. The fact that a large number of queries remain as far as the application of learning theory to behavior therapy is concerned, and indeed, as far as the establishment of learning theory itself is involved, cannot be an argument against the new paradigm. Quite the contrary; it is because this is a new paradigm that these problems remain for normal science to settle. Some of the advances made since the days of Breger and McGaugh will be recounted in this volume. A more detailed discussion of the point has been given by Eysenck (1976) in his chapter entitled "Behavior Therapy—Dogma or Applied Science?"

Among the critics of learning theory as a basis of behavior therapy, perhaps the most notable is Wolpe (1976a, b). He argues that behavior therapy is a synthetic construct, and is to be defined in terms of "principles and paradigms" rather than "learning theory." The case is argued in detail by Eaglen (1978), who concludes that

the development of theories . . . and their careful application to treatment programs is vital for the future development of behavior therapy, and it is only by insisting on a close link between therapy, theory development and research evidence that we can avoid the otherwise inevitable eclipse. (p. 128)

Wolpe's definition encounters the obvious danger that it may seem to encourage the "broad-band" eclecticism that he himself criticizes (Wolpe, 1976a, b).

It is difficult to see how it can seriously be argued that behavior therapists, in developing their methods, do not base themselves on principles of learning and conditioning theory. Wolpe (1958) himself clearly has done so in his development of desensitization therapy, and obviously Watson and Rayner (1920), in suggesting detailed methods of treatment to Jones (1924), had a clear theoretical rationale of Pavlovian extinction in mind. The work done under my direction at the Institute of

Psychiatry (Eysenck & Rachman, 1965; Rachman & Hodgson, 1980) has always been buttressed by reference to theoretical positions in learning theory, and most of the authors who have contributed original material in this field have clearly drawn on the same treasure house. It is not necessary to assert that they all make use of the same basic theories, or make identical deductions, but it would be difficult to find anyone who claimed to have derived his methods without benefit of prior experimental and theoretical work on the principles of learning and conditioning.

This is true even of those who have chosen eclecticism as a way of life. As Eysenck (1970a) has made clear, their applied work is characterized not by a rejection of theory as such, but rather by an overindulgence in reliance on heterogeneous theories that do not form any sensible kind of whole, and are often contradictory, partial, and difficult to integrate in a meaningful manner. Eysenck (1987) gave several examples of the development of methods of behavior therapy following on experimental laboratory work with animals, and the theories associated with that work. Critics seldom examine particular cases in order to try to demonstrate the alleged absence of reliance on theoretical formulations; they make wholesale suggestions without specifying the precise methods of therapy they have in mind as not being influenced by learning theory.

Altogether it seems that psychologists, possibly because they are often derided because their science, as William James suggested, was merely the "hope of a science," tend to take theories and their defects much more seriously than do hard scientists. Take as an example research in cryogenics. As Mendelsohn (1966) pointed out:

As was inevitable, ever since superconductivity was first discovered, many different theories with explanations have been proposed; roughly at the rate of 2 or 3 per annum, and for the better part of half a century. . . . [E]ventually Felix Bloch, who has done so much for our understanding of electrons in metals, announced an axiom of his own which ran: "every theory of super-conductivity can be proved wrong." And for a long time this axiom turned out to be the only correct one.

Yet in spite of this unpropitious state of affairs, theories of superconductivity have been used from the beginning to further practical ends, and the results of applied research have been used to disconfirm or improve existing theories.

Wolpe's (1976a, b) argument that because there is no one "modern learning theory," it is meaningless to define behavior therapy in terms of such a theory, and that instead treatment should be based on "principles and paradigms" is clearly one that hard scientists would not accept. We do have more than one theory, as is true of people working on cryogenics, and the correct way to use these is to make differential predictions from different theories, as far as application is concerned, and then study the results in order to choose between different theories. This is the way in which physics and chemistry have advanced so successfully over the last three centuries, and there is no reason to assume that psychology is positioned any differently.

But, it may be objected, is it not true that there are many different behavior therapies, rather than one single behavior therapy? The fact is not to be doubted, but its interpretation would seem to be somewhat different. The theory maintains that all cures of neurotic disorders are based on Pavlovian extinction, hence it makes sense to talk of behavior therapy. However, extinction can be produced along many different lines, but always involving the unreinforced exposure of the conditional

stimulus, either *in vivo*, or else in imagination. Among the best known procedures for producing extinction, we have modeling, desensitization, and flooding with response prevention, as well as many combinations and modification of these. Thus those who insist that there is one behavior therapy, and those who would prefer to talk about different "behavior therapies" are both right; there is one underlying principle on which all behavior (and other) therapies operate, but the application of this principle can take an infinite variety of forms.

CLAIMS OF COGNITIVE BEHAVIOR THERAPISTS

Cognitive psychologists, like Beck (1976), Mahoney (1974), and Meichenbaum (1977) have made strong and increasingly inclusive claims for what they sometimes term cognitive behavior therapy, with the stress more on the cognitive than on the behavioral side. Actually it is very difficult indeed to find any coherent account of theories, deductions, and experiments relevant to the claims made. Allport (1975) characterized the whole field of cognitive psychology in a rather unflattering summary. It is, he maintains, typified by

an uncritical, or a selective, or frankly cavalier attitude to experimental data; a pervasive atmosphere of special pleading; a curious parochialism in acknowledging even the existence of other workers, and other approaches, to the phenomena under discussion; interpretations of data relying on multiple, arbitrary choice-points; and underlying all else a near vacuum of theoretical structure within which to interrelate different sets of experimental results, or to direct the search for significant new phenomena.

M. W. Eysenck (1984), in his *Handbook of Cognitive Processes*, points out "the extremely diverse and sprawling nature of the current scene" (in cognitive psychology). And he goes on to say that

at least part of the reason for the growing army marching behind the banner of cognitive psychology is the increased vagueness with which the term is used. Virtually all those interested in perception, learning, memory, language, concept formation, problem solving, or thinking call themselves cognitive psychologists, despite the great diversity of experimental and theoretical approaches to be found in these various areas. (p. 1)

Eysenck finally characterizes cognitive psychology in terms of its "strong reaction against the facile approach of Behaviorism" (p. 2). This no doubt is true, but a reading of the criticism of behaviorism contributed by leading cognitive psychologists suggests that their criticisms are directed at the 1920 reflexological model, rather than the more up-to-date and much more formidable modern neobehaviorism presented, for instance, in Zuriff's (1985) book.

It is one of the sad features of this debate that it seems to be quite tangential to the real claims and issues. There can be no doubt about the real strength of modern neobehaviorism, yet sadly enough cognitivists disregard it completely, and direct their arguments at out-of-date beliefs, no longer held by behaviorists. Equally, behaviorists tend to play down the very real contributions made in many different fields by cognitivists, as outlined by M. W. Eysenck (1984). In so far as cognitive psychologists attempt to introduce cognitive concept into the behaviorist framework, they are simply following in the footsteps of Pavlov, who argued powerfully that

words can be both conditioned stimuli and conditioned responses, with the implication that cognitive events follow the same laws as behavioral events (Ullmann, 1981)

In talking about behavior therapy, we should of course bear in mind that different types of behavior may not covary in time in any precise manner. Lang (1970) endeavored to construct a three-system analysis of fear reactions, and later work by Hodgson and Rachman (1974) and by Rachman and Hodgson (1974) suggests that fear/anxiety reactions could best be regarded as consisting of at least three loosely coupled systems—subjective, behavioral, and physiological (see also Grey, Sartory, & Rachman, 1979). These authors have reviewed the evidence to show that one or more of these systems can be discordant at any particular time, and can change more rapidly or more slowly than the others in response to treatment (desynchrony). Some behaviorists might refuse to recognize the subjective reactions that make up one of the three systems as truly subject to scientific analysis, because of the well-known objection to introspective evidence of most behaviorists (Zuriff, 1985). Indeed, if it be true, as Nisbett and Wilson (1977) maintain, that subjects are sometimes unaware of the existence of a stimulus that importantly influences a response, are unaware of the existence of the response, and are unaware that the stimulus has affected the response, then indeed we might have to follow the restrictionist line of argument. However, as Nisbett and Wilson point out, accurate reports do occur when influential stimuli are salient and are plausible causes of the responses they produce; this would seem to cover the cases of fear/anxiety in neurotic patients, and the changes that take place during therapy.

It might be thought that the strong evidence regarding the primacy of affect (Rachman, 1981; Zajonc, 1980, 1984) would rule out subjective reports as important systems, but surely this is not so. Even though affect may precede cognition, and affective arousal may not always entail prior cognitive appraisal, this does not rule out the existence of such cognitive appraisal, and its importance for the patient.

Recognition of the desynchrony of affect has led to an interesting paradox pointed out by Bandura (1977). As he states,

on the one hand, explanations of change processes are becoming more cognitive. On the other hand, it is performance based treatments that are proving most powerful in effecting psychological changes. Regardless of the method involved, the treatments implemented through actual performance achieve results consistently superior to those in which fears are eliminated through cognitive representations of threats. (p. 78)

Bandura (1977), goes on to argue that

the apparent divergence of theory and practice will be reconciled by recognizing that change is mediated through cognitive processes, but the cognitive events are induced and altered most readily by experiences of mastery arising from successful performance. (p. 193)

Rachman and Hodgson (1974) draw the inference that different methods of behavior therapy may be appropriate, depending on which of the three systems is most deviant, and make appropriate suggestions in this respect. They are more willing than most behaviorists to accept the subjective, introspective, mental type of fear reaction as equally important with the behavioral and physiological types of response, and of course cognitive psychologists not only accept this, but emphasize cognitive components, sometimes to the exclusion of physiological and behavioral ones. The

growing stress on cognitive components, mentioned by Bandura, makes it necessary to devote some consideration to their claims.

We may now turn to a consideration of the specific arguments advanced by cognitivists in the field of behavior therapy. The position taken here is essentially that of Wolpe (1978), who argued that cognition is also behavior and is subject to the same law of inevitability as other behavior. Accepting the Lang-Rachman-Hodgson theory of desynchrony, he argues that "overcoming the unadaptive learned habits typified by neurotic behavior requires cognitive, autonomic and motor relearning, according to the indications of behavior analysis" (p. 437). If we can embody cognitive processes within a behavioristic framework, what then is the major contribution of cognitive psychologists?

Marzillier (1980) pointed to three major usages made by cognitive therapists: cognitive events, cognitive processes, and cognitive structures.

Cognitive events have been readily assimilated into behavior therapy, and, as Beck pointed out, they have been there from the beginning. What has emerged has been the increasing interest in cognitive events as dependent variables, the focus on client's thoughts and images in relationship to their emotional problems. This is an area that cognitive therapy has pioneered, and its techniques and practices are of value to behavior therapists. Much less attention has been paid to cognitive processes in behavior therapy. However, it is evident from recent developments in behavior therapy, that therapists are beginning to focus directly on deficiencies in cognitive processing, using cognitive restructuring and problem-solving methods as part of the behavior approach. Particular attention is drawn to the implications of cognitive appraisal which can be seen as stressing the meaning of events and behavior. In behavioral analysis a concern for meaning should provide greater breadth and sensitivity in the process of assessment. Finally, cognitive therapists have stressed the need to consider long-term fundamental cognitive change, as a goal of therapy. The term "cognitive structures" has been used but as yet it lacks precise meaning. It is possible for behavior therapists to consider cognitive structures, such as beliefs or attitudes, as these can be linked closely to observable behavior. However, the value of so doing remains to be established. (p. 256)

Are these alleged contributions to behavior therapy real, or are they merely promissory notes drawn against a nonexistent account? Latimer and Sweet (1984) gave a critical review of the evidence concerning cognitive versus behavioral procedures in cognitive-behavior therapy. They mention the increasing emphasis on cognition in psychology and behavior therapy during the past decade, and address the question of whether cognitive therapy is an evolutionary or revolutionary development from behavior therapy. They also critically evaluate the evidence for the efficacy of procedures specific to cognitive therapy. Their conclusions are worth quoting in full.

Cognitive therapy is an evolutionary rather than a revolutionary development in the field of behaviour therapy. It is unique only in its greater emphasis on one class of behavior—cognitions. Several innovative therapeutic methods have been spawned as a result of this shift and emphasis, but these have not been demonstrated to be efficacious in the treatment of clinical populations. Cognitive therapy as actually practiced usually involves a variety of methods including behavioral procedures of established efficacy. Most of the claims made in support of cognitive therapy are based on studies employing these cognitive-behavioral methods. It remains to be demonstrated either that the new cognitive therapy procedures make a significant contribution to therapeutic outcome or that existing behavioral methods are rendered more effective when conceptualized in cognitive terms. The widespread adoption

of cognitive treatment procedures is unwarranted on the basis of existing outcome data involving clinical populations. (p. 21)

This paragraph highlights the problem that is posed by the claims made by cognitivists. The main claim of behavior therapy to a higher scientific status than psychotherapy has always been its willingness to attempt to prove its assertions by actual clinical experiments, and to employ empirical comparisons between different types of treatment in order to establish the superiority of one over the others. Cognitive psychologists have reverted to the older practices of psychoanalysts and psychotherapists, all making claims without furnishing proof that these claims are actually justified. By using mixed methods of treatment they make it impossible to distinguish the contribution of cognitive and behavioral variables. Until and unless they bring forward actual experimental proof of the superiority of their methods, it is impossible to concede these claims. By presenting these claims as a "paradigm shift" they suggest a successful revolution in our conception of neurosis and therapy, but there is really no justification for this claim (Eysenck, 1987). Revolutions in science are based on demonstrated effectiveness, not on speculation and theoretical argument unsupported by empirical data. There has been a paradigm shift in this field, but it has been from psychotherapy to behavior therapy. The possibility exists that the theoretical advances noted by Marzillier will result in an actual demonstrable improvement in the rate of recovery of neurotic patients, but until this has been clearly substantiated it would be premature to jump on this particular bandwagon.

'It may be useful, at the end of this section, to note some of the ways neobehaviorism (or what Davey, 1983, a,b, calls dialectical behaviorism) attempts to cope with the complexities of human behavior that are not normally covered by the older forms of behaviorism. First and foremost, as we have seen, it uses words and language as part of a conditioning system (Platonov, 1959; Staats, 1964, 1968). In the second place, we have the work of Levey and Martin (1983) and Martin and Levey (1985) to demonstrate the existence of evaluative conditioning as a process that uses the principles of conditioning in a specifically human context. Third, we have the insistence on the importance of central representations in the conditioning process.

Mackintosh (1984) makes it quite clear how views have changed in recent years as far as learning theory is concerned.

The view of conditioning as the establishment of new reflexes or the strengthening of S-R connections, a view which dominated Western learning theory for half a century, has gradually given way to a view of conditioning as the acquisition of knowledge about the relationship between events in an animal's environment, knowledge which may not be immediately apparent in any change in behavior at all. When a CS is regularly followed by a reinforcer, animals can be said to learn that the CS signals the reinforcer. This is achieved by the establishment of an association between some central representations of the two. From studies that have altered the value of a reinforcer after conditioning, it is apparent that the representation of the reinforcer associated with the CS must, in at least some cases, itself be available for modification when their value is manipulated. (p. 56)

Such "central representations" are of course cognitive processes as properly defined, and hence modern learning theory incorporates cognitive processes in a most explicit fashion; and it is not clear why cognitive theorists should claim exclusive patronage of such processes, or why they should declare that modern learning theory is incomplete

because it does not take them into account, when clearly it does (Rescorla, 1972). A more detailed discussion of this whole problem is given by Eysenck (1987).

The S-S analysis of conditioning can very easily be integrated with an information-processing paradigm, which reflects the more empirical contributions that cognitive psychology has to make to learning theory (Kanfer & Hagerman, 1985). Reiss (1980) and Bootzin (1985) outline such a theory following Wagner and Rescorla's (1972) information model. These developments are implicit in Tolman's (1948) view of learning theory, and do not present an alternative view to theories of cognition. Altogether, those who oppose cognitive to conditioning theory would seem to commit the logical fallacy technically known as the *unacceptable entymeme*. It presents an argument with one of its stages understood rather than stated, the understood premise being that conditioning theories are of the Watsonian S-R type; this premise is clearly erroneous.

A NEW CONDITIONING MODEL OF NEUROSIS

Watson's conditioning theory has encountered many criticisms, and at first sight these seem fatal to it, certainly in its original form. Some of these will be considered in a later chapter by the present author; others will be considered here. Eysenck (1979, 1982a) has pointed out that Watson's theory (like Freud's) is based on the occurrence of a traumatic fear-producing event constituting the Unconditional Stimulus (UCS), which is followed by fear/pain responses that constitute the (UCR) Unconditional Response. Neutral stimuli accidentally present at the time will become conditioned through contiguity, thus being made into CSs that from then on will evoke CRs similar in nature to the UCR, that is, feelings of fear and pain. These CRs continue indefinitely, unless deconditioned along the lines discussed by Watson and Rayner (1920), and exemplified in the work of Mary Cover Jones (1924). This account raises the following problems.

The first problem is a clinical one. War neuroses often do begin with a traumatic event, such as the person in question being buried alive by an explosion, or coming into contact with death or mutilation of friends and colleagues. However, in civilian neuroses such events are very rare, and in the majority of cases the initiating event is not excessively traumatic, and does not produce an immediate, strong CR. Rather, there appears to be an insidious increase in the anxiety produced by the CS that may take years, or even decades before a full-blown phobia becomes apparent, or a clinical state of anxiety is reached. This is the major clinical objection to the theory.

From an experimental point of view, a second objection is the simple one that on this account extinction should set in almost immediately, making impossible the development of any long-lasting neurosis. Whatever the CS may be, the subject is likely to encounter it quite frequently and without attending reinforcement. This should produce relatively quick extinction of the CR. Let us consider a person suffering from a cat phobia; he or she is likely to encounter cats in nonthreatening situations quite frequently, and each such encounter should foster extinction. The phobia should thus quite soon disappear. The fact that this does not seem to happen is a powerful argument against Watson's theory (Kimmel, 1975).

A third point of importance is that in ordinary Pavlovian conditioning there is

no way in which the CR could be stronger than the UCR. Yet if we look at clinical cases, as mentioned earlier, the initiating conditioning experience often leads to UCRs and CRs that are rather mild; it is only after the insidious development of the neurosis has taken place that the CRs become so strong as to constitute an actual mental illness. Hence in these quite typical cases of neuroses and phobias, the CR becomes much stronger than the original UCR; on ordinary Pavlovian principles this would seem to be impossible.

What these three objections have in common, of course, is a reference to the development of the CR over time, when the subject is exposed a number of times to the CS only, that is, to the CS without simultaneous reinforcement. Classical conditioning theory would expect extinction under these conditions, but what happens in the case of the development of a neurotic illness seems to be the opposite, that is, an incrementation of the CR. To explain this anomaly, Eysenck followed up Grant's (1964) suggestion that there was an important distinction between Pavlovian A and Pavlovian B conditioning, and proposed that the consequences of this distinction are important in regard to extinction. (Eysenck, 1967, 1968, 1976, 1979, 1980, 1982a, b)

Pavlovian A conditioning is exemplified by the textbook example of classical conditioning, that is, salivation on the part of the dog to the sound of a bell that had been repeatedly presented shortly before food was given to the hungry dog. Of the many UCRs presented to the dog (approach to the food, ingestion, etc.), Pavlov chose only to measure one, namely buccal salivation. As Zener (1937) pointed out, it is noteworthy that the CR did not include approach to and attempts to feed upon the bell or other source of the CS. Any approach and reorientation movements were directed to the food source, showing that the CS does not substitute for the UCS, as S-R theorists have often stated. Pavlov maintained that the CS serves as a signal that the food is about to be presented, and this position is also taken by S-S theorists. This approach is now almost universally recognized as being more in line with the facts than the old-fashioned S-R approach (Mackintosh, 1984).

Pavlovian B conditioning is directly linked by Grant (1964) to the Watson and Rayner (1920) experiment, but as he points out, Pavlov has priority. A reference experiment for Pavlovian B conditioning could be that in which an animal is given repeated injections of morphine. The UCR in this case involves severe nausea, profuse secretion of saliva, vomiting, and then profound sleep. After repeated daily injections, Pavlov's dogs were found to show severe nausea and profuse secretion of saliva at the first touch of the experimenter (Pavlov, 1927, p. 35-36).

The major differences between Pavlovian A and B conditioning relate to drive, and degree of similarity between CR and UCR. In Pavlovian A conditioning, no learning takes place unless the subject is in a suitable state of drive, such as hunger in the case of salivary conditioning in dogs. In the case of Pavlovian B conditioning, the UCS provides the drive or motivation. In Watson's theory the UCS clearly provides the drive, making his a case of Pavlovian B conditioning.

In Pavlovian B conditioning, the UCS elicits the complete UCR, whereas in Pavlovian A conditioning the organism emits the UCR of approaching and ingesting the food. Thus in Pavlovian B conditioning the CS appears to act as a (partial) substitute for the UCS, which is not true of Pavlovian A conditioning. Expressed in different terms, we may say that in Pavlovian A conditioning typically the CR and

the UCR are different (salivation as opposed to approach to and ingestion of food), whereas in Pavlovian B conditioning they are similar or identical (nausea, profuse secretion of saliva, vomiting). As Grant points out, many components of the UCR in Pavlovian conditioning "are readily seen as components of the CS which will be evoked by the preparations of the injection after repeated daily morphine injections." (p. 5). A great deal of interoceptive conditioning (Bykov, 1957) and autonomic conditioning (Kimble, 1961) appears to follow the Pavlovian B paradigm.

These differences between Pavlovian A and Pavlovian B conditioning can be used to argue that the consequences of CS-only presentations may be quite different in the two paradigms. (Eysenck, 1976). In Pavlovian A conditioning, it is meaningful for both the subject and experimenter to talk about CS-only presentation as the presentation of the CS that is not followed by the UCS. However, in Pavlovian B conditioning this is difficult to accomplish because the CR, which follows the CS, is for all purposes identical with the UCR. Consequently, the phrase *CS-only presentation* is meaningful for the experimenter, who controls the presentation of the UCS, but not for the subject, who experiences the CR as identical with the UCR. In Pavlovian B conditioning, if it be true that the CS-only condition is not necessarily fulfilled (as far as the subject of the experiment is concerned), then it would seem to follow that the ordinary laws of extinction might not always apply. Although the experimenter has arranged the contingencies in such a way that CS is not followed by UCS, under certain conditions (to be specified later) the CR itself might act as a reinforcement equivalent to the UCR, thus producing not extinction but an increment in the strength of the CR. This incrementation has been called *incubation* and has led to a revised conditioning theory of neurosis (Eysenck, 1968, 1976).

There has been much discussion of the incubation phenomenon, and the large body of research that supports it (Eysenck, 1976, 1979, 1982a); there is no space to review the evidence again here. Incubation is a process that is theoretically intelligible in terms of Pavlovian B conditioning, and experimentally verified by many animals and a few human experiments. We also have both theoretical and practical evidence concerning some of the variables that make for incubation rather than extinction, such as strength of the UCR and CR, duration of exposure of the CS-only, personality, etc. (Eysenck, 1982b).

The general form of the theory of incubation and extinction in neurotic fear reduction is shown in diagrammatic form in Figure 3. It shows on the ordinate the strength of the CR, and on the abscissa the duration of CS-only exposure. Curve A illustrates the decline in fear/anxiety with duration of CS exposure; there is ample evidence from the animal and particularly from the human field (Rachman & Hodgson, 1980, Figure 14.1) to support the general decline over time of the fear/anxiety reaction. The theory states that on this curve there is a critical point. If CS-only exposure stops before this point is reached, that is, while the strength of the CR is above this level, incubation will result. If at termination of CS-only exposure the strength of the CR is below this critical point, extinction will result. Thus duration of exposure is a critical element in deciding whether incubation or extinction is to result from treatment or experiment, and there is much evidence from the clinical field to support this view (Eysenck, 1982a, 1983, 1986; Eysenck & Beech, 1971).

If CS-only exposure is continued long enough to provide an increment of extinction, Curve A will be lowered on the next occasion, as it is indicated by Curve B,

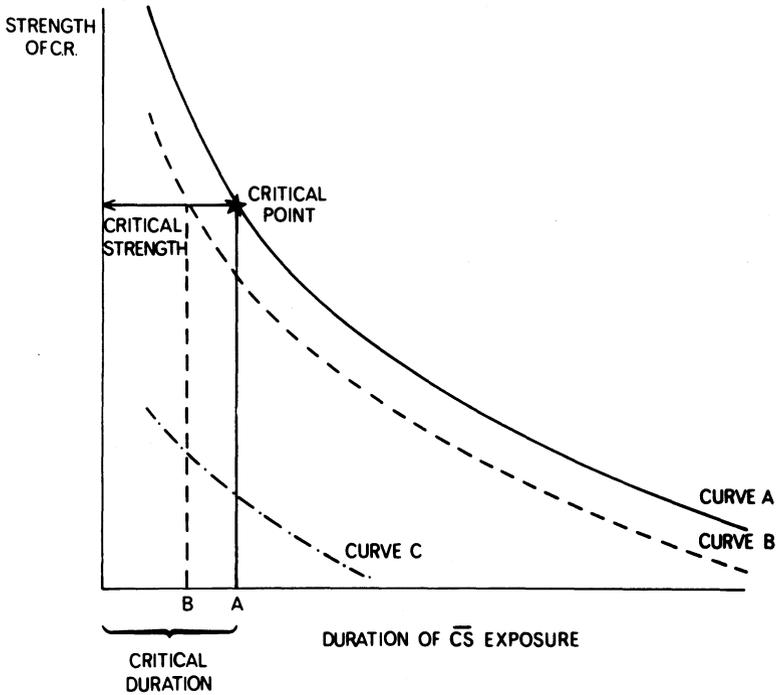


FIGURE 3. Diagram of incubation versus extinction theory.

and subsequent increments of extinction will reduce the whole curve below the critical point, as in Curve C. Curve A indicates a typical sequence of events when flooding with response prevention is used as a therapeutic technique; Curve C indicates the level at which desensitization and modeling proceed.

Strength of the CR and duration of CS-only exposure are not the only critical variables; as we shall see later, personality (and the concentration of peptides and hormones that control both personality and fear/anxiety reactions) also play an important part. Note that the theory is also relevant to the acquisition of fear/anxiety responses; if the original CR exceeds the critical point, then incubation will occur and the final CR will be stronger than the original UCR, an event not contemplated in Pavlov's original theory, but clearly apparent in experimental animal studies, as well as characteristic of the development of human neuroses (Eysenck, 1982a, 1986.)

THE NEUROBIOLOGY OF INCUBATION OF FEAR/ANXIETY*

We can trace the variegated events of extinction and incubation a little further into the biological realm by considering individual differences in levels of neurohormones. The hypothesis developed by Eysenck and Kelley (1987) largely stems from 35 years of animal research that has shown that neurohormones can have a profound modulating influence on resistance to extinction. Considerable experimental and

*This section is paraphrased from a more detailed account by Eysenck and Kelley (1987).

clinical work with humans is consistent with the possibility that it is individual differences in these hormones that mediate the persistence characteristic of disorders such as phobias, and the absence of persistence typical of depression. In relation to the incubation concept, Eysenck and Kelley argue that individual differences in levels of peptides, such as ACTH, allow the fear-producing CSs to increase dramatically in excitatory strength, or to decrease and extinguish, depending on the hormone and the duration of CS exposure. The literature suggests that hormonal mediation of incubation is a reliable phenomenon. At the level of psychological processes, it is suggested that incubation occurs by hormones influencing mechanisms of attention so as to produce changes in CS associability or in the absolute capacity of a CS to have inhibitory or excitatory strength. At the level of psychological treatment of neurosis, this model predicts that an intervention strategy involving both hormones and conditioning may have more impact than manipulation of only one of these factors.

Apart from peripheral endocrine functions, hormones are present in the CNS and affect emotions by the modulation of activity in the limbic system. Patients with panic attacks, for instance, have limbic abnormalities (Reiman, Raichle, Butler, Herscovitch, & Robins, 1984), and it is known that the behavioral effects of hormones are dependent on the integrity of limbic structures (de Wied & Jolles, 1982; Van Wimersma Greidanus *et al.*, 1983). In addition, hormone-induced changes in hippocampal theta occur that show some correspondence with anxiety-related behavioral outcomes (Gray, 1982; Urban, 1984; Urban & de Wied, 1978). This modulation of limbic activity is the balanced outcome of many hormones. The hormones have precedence in the course of evolution, and are of at least equal importance as the better-studied neurotransmitters (Iverson, 1984; Krieger, 1983; Le Roith, Shiloach, & Roth, 1982).

It may be useful to begin our discussion with a mention of some of the findings that suggest a relationship between neurohormones and neurosis. Redmond (1981) and Hall (1979) have demonstrated that there are similarities in the symptoms of anxiety neuroses and withdrawal from opiate addiction, an observation consistent with the finding that there is a strong negative correlation (-0.67) between levels of trait neuroticism and opioid peptides in the cerebrospinal fluid (Öst & Hugdahl, 1981). This correlation is even higher (-0.91) when a measure of state anxiety is employed; from this relationship, and the well-established relationship between analgesia and CNS opioids it is possible to argue that low levels of opioids in the brain of neurotics may make them more susceptible to incubation effects. This point, and the literature relevant to it, are both dealt with in much more detail by Eysenck and Kelley (1987).

Another important peptide hormone is adrenocorticotrophin (ACTH). Whereas opioids dampen neuronal excitation, cholinergic and noradrenergic turnover rates, and behavioral performance in aversive conditioning, ACTH has the opposite effect (Bertolini and Gessa, 1981; Chorney & Redmond, 1983; Markey & Sze, 1984; Redmond, and Huang, 1979; Redmond and Krystal, 1984). In the ACTH-mediated incubation effects, which we will discuss in detail shortly, the opioids have a competitive affinity with ACTH for the same receptors. If we block these receptors with the opioid antagonist, naloxone, ACTH (and also vasopressin) loses its capacity to

induce incubation effects or prolong extinction (Concannon, Riccio, Maloney, & McKelvey, 1980; Concannon, Riccio, & McKelvey, 1980; De Vito & Brush, 1984). This reciprocal relationship between ACTH and the opioids, in conjunction with the negative correlation between CSF opioids and anxiety, suggests that ACTH may play an active role in the occurrence of incubation effects.

A tie between experimentally produced changes in emotionality (defecation) as a trait, and the capacity for stress-induced changes in ACTH levels has been recently demonstrated by Armario, Castellanos, & Balasch (1984). This observation can be combined with the findings of Morley (1977), who showed that emotional animals are more likely to show incubation effects. This is also consistent with the suggestion of Eysenck (1979, 1982a) that incubation effects are likely to be stronger in subjects high on neuroticism (N) and introversion (I). The reasons for this suggestion can be deduced from the nature of these two major personality dimensions (Eysenck & Eysenck, 1985), and need not be detailed here. We will now turn to direct experimental evidence that hormones such as ACTH can modulate incubation.

In a series of aversive conditioning studies by Riccio and his students, ACTH or epinephrine injections (which increases ACTH in the rat) were given to rats prior to a one-minute presentation of the CS during a forced-exposure trial following acquisition training. This procedure repeatedly resulted in a large permanent increase in fear of the CS when animals were tested 24 hours later for resistance to extinction without an injection. Mere presentation of the CS or elevation of ACTH levels alone did not produce such effects (Haroutunian & Riccio, 1977, 1979). Kelley (in press) provided an additional control. In this experiment rats were first given three .5 ma foot shocks during two direct placements on the black side of a shuttle-box with a closed guillotine door, and never shocked during two placements on the white side. In the second phase, the different groups of rats were reexposed to the black side and given either (a) a prior .02 mg injection of epinephrine or (b) saline, or (c) an epinephrine injection 5 hours later. The latency to cross from the white to the black side 24 hours later was found to be several-fold longer in the groups given an epinephrine injection shortly before reexposure. The findings thus demonstrate that contiguity between the presence of the fear cue and high levels of hormones is required to produce incubation effects in the rat. The importance of this contiguity has also been demonstrated by other investigators (Righter, Elbertse, & van Riezen, 1975; Weinberger, Gold, & Sternberg, 1984). Whereas ACTH released by acute exogenous injections is one possible explanation of this, it is also possible that epinephrine itself is important (Borrell, De Kloet, Versteeg, & Bohus, 1983; McGaugh, 1983).

The capacity for ACTH to produce incubation effects is supported by an extensive body of evidence from many laboratories showing that ACTH will enhance resistance to extinction. This occurs with a variety of aversive conditioning procedures and with ACTH (4–10), which has CNS but no peripheral endocrine properties (de Wied & Jolles, 1982). There is also evidence that physical levels of ACTH can have a modulatory effect on extinction (Bohas, Endrocrozi, Kissak, Fekete, & de Weid, 1970; Pagano & Lovely, 1972; Van Wimersma Gredanus *et al.*, 1977, 1983). Although these properties of ACTH have also been observed in Pavlovian A conditioning, they are more readily observed with Pavlovian B conditioning. Consistent with this is the well-established observation that sexual behavior is the other motivational system

where ACTH has a robust effect (Bertolini, Fratta, Eena, Munladr, & Serra, 1981, 1984; de Wied & Jolles, 1982). Eysenck (1982a) suggested that sexual drives are the equivalent on the appetitive side to anxiety on the aversive side for the production of incubation effects (i.e., are mediated by Pavlovian B conditioning).

These findings are consistent with other open-field research suggesting that injections of ACTH or its releasing factor (CRF) may be “anxiogenic” in rats (Britton & Britton, 1981, 1982; File & Vellucci, 1978). Some support for this hypothesis also comes from studies on the effect of CRF in rhesus monkeys (Kalin *et al.*, 1983a, b); however, these “anxiogenic” properties of CRF and ACTH are only seen in situations that are already fearful. For instance, in the study by Haroutunian and Riccio (1979), exposure to one side of a novel shuttle-box contiguous with an ACTH injection was not itself sufficient to produce later spatial avoidance of that side of the apparatus; thus it would appear that an ACTH injection is not, by itself, an aversive UCS. Considerable evidence suggests that the action of ACTH is on the CS, not the UCS or UCR. When ACTH levels are increased by adrenalectomy, the immediate behavioral responses to foot shock (flinch, jerk, vocalization) are not increased (Borrell *et al.*, 1983). In addition, although reduced open-field ambulation is sometimes observed after adrenalectomy, this is not influenced by injections of dexamethasone that should reduce the ACTH levels. Similarly, effects on exploratory behavior are not reliably found after injections with ACTH 4–10 (Bohus *et al.*, 1982). The open-field apparatus has been shown to be a potent releaser of fear (Blanchard, Kelley, & Blanchard, 1974) but that might be dependent upon the strain of rats used, which would account for some of the ambiguity in the open-field findings with ACTH (Eysenck & Broadhurst, 1964). Finally, the results of experiments with humans also suggest that injections of CRF or ACTH are not themselves anxiogenic (Beckwith & Sandman, 1978, 1982; Gold *et al.*, 1984). In contrast, the anxiogenic properties of the ACTH in rats are readily observed when a CS for fear is present; then, as we have just seen, ACTH enhances the excitatory properties of the cue.

Another possible explanation is that incubation effects are mediated by selective attention: ACTH and other neuromodulators of anxiety may enhance the capacity of a CS to show an increment in the level of excitatory strength by influencing its associability. Unfortunately, the effects of ACTH and other peptides on blocking and overshadowing—indexes of selective attention in rats—have not as yet been investigated (Mackintosh, 1984); thus direct evidence for this possibility awaits testing. Nevertheless, Beckwith and Sandman (1978, 1982), using reversal learning and intra- and extradimensional shift experiments, have argued that ACTH influences selective attention. These older behavioral assays for selective attention, however, are subject to alternative interpretations (Mackintosh, 1974; Sutherland & Mackintosh, 1971).

This is not the place to continue the detailed discussion of the action of hormones, which could be extended to vasopressin and cortisol (Eysenck & Kelley, 1987), all of which have been shown to influence conditioning and extinction behavior. The main point of this section has been merely to point to the importance of extending the field of search for causes or mechanisms to that of peptides and other neurohormonal factors that have been clearly related in numerous experimental studies to fearful, neurotic types of behavior, conditioning, and learning, and also incubation.

It seems likely that a better understanding of the mediation of Pavlovian conditioning and extinction can be obtained by looking more deeply into the biology of the organism, rather than treating it merely as a black box, as behaviorists are wont to do (Zuriff, 1985.)

IS EXPOSURE A NECESSARY CONDITION FOR FEAR REDUCTION?

The essence of Watson's theory, both in its original form, and as amended by the writer, is the notion that neuroses are the product of Pavlovian conditioning, and cures are mediated by Pavlovian extinction. We have noted in previous sections that the concept of Pavlovian conditioning has changed very much from the original simple connection between primitive sensory impressions and muscle twitches, to the much more sophisticated type of S-S conditioning that forms the basis of neobehaviorism and of dialectical behaviorism (Davey, 1983a, b; Mackintosh, 1984.) In this section we will attempt to look at some of the unifying properties of the theory, and go on to confront the question raised by de Silva and Rachman (1981) of whether exposure is really a necessary condition for fear reduction. We will also consider some of theoretical implications raised by this question for the relation between cognitive and behavioral theories.

Eysenck (1980, 1983, 1985) has suggested that this (modified) conditioning theory of neurosis can explain the major facts that are known about the treatment of neurotic illnesses. These may be summarized as follows. (a) Spontaneous remission, that is, fear reduction without any form of psychiatric treatment, is a fairly regular and very important factor in the improvement or cure of people suffering from neurotic illnesses. (b) Placebo treatment is a very successful means of securing fear/anxiety reduction in neurotic patients. (c) Nonspecific psychotherapeutic intervention is as successful as placebo treatment, and possibly slightly more so than spontaneous remission, in securing reduction of fear/anxiety reactions in neurotic patients. (The term *nonspecific* is meant to denote the fact that regardless of the theoretical basis of the therapies in question, they are equally successful, suggesting the lack of relevance of the specific theories on which they may be based.) (d) Psychoanalysis on the whole is no more successful than all other methods of psychotherapy, and may be less so. (e) Psychoanalysis specifically has been found to have frequent negative treatment effects, that is, it increases rather than reduces fear/anxiety reactions (Strupp *et al.*, 1977.)

It is possible to explain all these effects in terms of exposure to the unreinforced CS. To begin with spontaneous remission, it is well known that people suffering from neurotic disorders, but unable to obtain psychiatric treatment, will seek out various ways of alleviating their distress, usually by discussing their problems with parents, priests, friends, or other people whose friendly advice they feel able to count upon. In such discussions they inevitably bring up the problems which confront them, and discuss in detail, often in some sort of hierarchical order, the fears and anxieties, and their causes, that they have encountered. Thus the relevant CSs, in the weakened form demanded by desensitization theory, are encountered in a relaxed setting in which a friendly listener provides additional relaxation. The presence of a friendly

listener is vital according to Zillmann's (1978, 1979, 1984) three-factor theory of emotion, and in particular the excitation-transfer paradigm to which it has given rise. This paradigm applies to potentially unrelated successive emotional reactions, and to emotional reactions solicited by simultaneously present, yet potentially unrelated stimuli.

In the former case, the paradigm projects the intensification of any emotional reaction that is evoked during the presence of residual sympathetic excitation from antecedent reactions—with some specifiable exceptions. In the latter, it projects the intensification of any emotional reaction by sympathetic excitation due to stimuli other than those that elicited the emotional reaction proper. The paradigm is applicable to all emotional reactions associated with sympathetic dominance in their excitatory component. (Zillmann, 1984, p. 148)

Zillmann's three-factor theory distinguishes between the dispositional, excitatory, and experiential components of emotional behavior, in which the excitatory activity of emotions that are characterized by sympathetic dominance in the autonomic nervous system is largely nonspecific, and hence capable of being subject to Zillmann's additive law.

What is posited, then, is that the relaxing effects of the presence of the friendly observer/listener summates *negatively* with the relatively slight sympathetic arousal produced in the "patient" by his evocation of the feared material, and thus assists in the general desensitization that is the outcome of this mode of exposure. It thus reduces the strength of the CR, as in Figure 3, making extinction more likely, and incubation less so. Clearly the effects will be much less clear-cut and marked for spontaneous remission than it is for properly planned desensitization at the hands of an experienced behavior therapist, because the contingencies are not planned, and are only accidentally likely to be optimal; nevertheless, the general combination of stimuli is similar, even if their sequence is nonoptimal, and consequently reduction in fear/anxiety behavior is to be expected.

Much the same is true of placebo treatments in so far as these usually implicate the evocation of material giving rise to few anxiety reactions in the patients, under relatively relaxing conditions. This is also true of most psychotherapies, where the presence of a friendly therapist, evoking fear/anxiety related material, is very likely to lead to a general situation the effects of which should be similar to desensitization. Thus the theory may account for spontaneous remission, placebo effects, and also the nonspecific effects of the various psychotherapies.

Psychoanalysis presents a particular difficulty because of the established fact that it often leads to negative emotional reactions, that is, to what we would call incubation rather than extinction of fear/anxiety reactions (Strupp *et al.*, 1977). The reasons for this are probably to be found in the well-known fact that psychoanalysts are instructed not to act in such a way as to appear helpful, warm, relaxing, and generally friendly, but rather are expected to preserve a remote and neutral presence, and to act only in an interpretative rather than an advisory manner. If we again use Zillmann's excitation-transfer paradigm, this failure to provide a helpful, relaxing atmosphere would increase the general level of fear/anxiety of the patient, thus leading to incubation rather than extinction of anxiety (Eysenck, 1982a). Individual accounts of these effects strongly suggest the accuracy of this view (Sutherland, 1976; York, 1966). A simple exposure theory therefore seems to be capable, in combination with

Zillmann's excitation-transfer theory, to explain the observed phenomena to a reasonable degree.

We must now turn to the question of whether exposure is a necessary, as well as a sufficient condition for fear/anxiety reduction. De Silva and Rachman (1981) defined the term *exposure* as

the exposure of the subject to the fear-evoking stimulus, either in real life (*in vivo* exposure) or in fantasy (imaginal exposure). Imaginal exposure . . . consists in planned, sustained and repetitive evocations of images/image sequences of the stimulæ in question. Mere thoughts or fleeting images do not constitute imaginal exposure in this sense. Indeed, it is worth noting that subjects who lack the ability to conjure up and maintain detailed and vivid imagery are usually excluded from therapies involving imaginal exposure. (p. 227)

(See Wolpe, 1973, pp. 136–137, Rimm & Masters, 1979, pp. 45–46.)

De Silva and Rachman (1981) argue that

while in many circumstances exposure may be a *sufficient* condition for fear-reduction, there is no good reason to suppose that exposure is a *necessary* condition for success. Fear reduction that takes place in the absence of such exposures undermines the assumption that exposure is a *necessary* condition. (p. 22)

De Silva and Rachman admit that

all of the examples [they give] are open to criticisms of one sort or another and we shall draw attention to these, but remain confident that in due course when the controlled experimental analyses are completed, the results will confirm our claim that fear can be reduced even in the absence of exposure. (p. 228)

This quotation from de Silva and Rachman makes it clear that there is no proper experimental evidence for their claim, but merely anecdotal evidence, and it will be shown that even that is not fatal to the interpretation of exposure as a necessary condition for fear/reduction.

The argument begins with a reference to Rachman's (1978) suggestion that there are at least three pathways to the acquisition of fear. The three pathways are (a) direct aversive experiences; (b) vicarious acquisition of fear;* (c) the fears that are induced by the transmission of information. De Silva and Rachman attempt to use the same arguments that were brought to bear in analyzing the acquisition of fear to the question of reduction.

It is argued that fears can be reduced by direct experiences (such as desensitization), by vicarious exposure to the therapeutic model, and—most important for the present argument—by the transmission of information. It seems to us to be indisputable that a person's fear reactions can be weakened or even eliminated by giving him the information that the fear stimulus or the surrounding circumstances are not dangerous. (p. 228)

De Silva and Rachman refer to "studies in which fearful subjects were successfully treated with cognitive-based techniques that did not include contact with the fear stimuli" (e.g., Meichenbaum, Gilmore, & Fedoravicius, 1971; Weiss, Nelson, & Odom, 1975; Weissberg, 1977). The same point is made by Bandura (1977), implicit

*Vicarious acquisition of fear, and its complement in therapy, modeling, are not necessarily inexplicable in terms of conditioning theory, as Baer and Teguchi (1985) have shown.

in whose theory is the denial of the claim that exposure is a necessary condition for behavior change such as reduction in fear.

I think it is essential to posit the existence of two continua which are involved in fear/anxiety. The first continuum ranges from the postulation of reasonable as opposed to unreasonable fears. It is reasonable to be afraid of torture, disease, or injury, in circumstances that are likely to result in any of these consequences. It is unreasonable to be afraid of spiders (in countries where there are no poisonous ones), of contamination (when quite innocent objects are being touched), or of many other typically neurotic fear/producing stimuli. These are not two classes of stimuli, because intermediate ones can be adduced. To what extent is fear of an atomic war realistic, to what extent neurotic? Fear of dogs, cats, and even squirrels can to some degree be justified in Europe where these animals may be infected with rabies.

Again, fears may be acquired through cognitive learning (as when we are told that another person is suffering from an infectious disease, or that we can be burnt by fire, or that lions and tigers are dangerous animals), or fears can be conditioned through Pavlovian B conditioning. This too is a continuum rather than an either/or type of classification, because cognitive and conditioning methods of fear acquisition may be active in any particular instance. Thus the obsessive-compulsive patient's fear of contamination is partly due to what he has been taught cognitively about the danger of bacteria, and is partly acquired through a process of conditioning. Note the demonstration by Öhman, Dimberg, and Öst (1985) that prepared CSs are very difficult to extinguish by cognitive means, as compared with nonprepared CSs.

As Figure 4 shows, we thus have four categories of fears. Some rational fears are learned, as are some irrational fears; indeed, it is likely that there is some contribution by cognitive learning in all fears. Similarly, it seems likely that even rational fears may have some slight, and often a considerable basis in Pavlovian conditioning. That this must be true is obvious from a consideration that the very data given

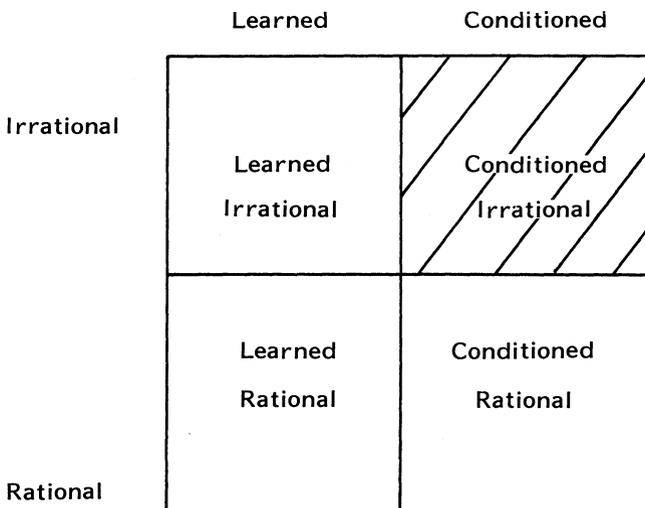


FIGURE 4. Different combinations of learned and conditioned, irrational and rational fears.

us by perception are acquired and organized in terms of conditioning (Taylor, 1962).

In this two-dimensional framework, neurosis is primarily concerned with conditioned irrational fears and anxieties, but the fact that we are dealing with continua means that we will hardly ever encounter a pure example, suggesting the possibility that part of these fears may be acquired through learning, and hence in part extinguished through unlearning. In that sense De Silva and Rachman are undoubtedly correct, but this does not detract in any way from the theory here maintained, namely that specifically neurotic fears are reduced only through exposure. Consider an example. This concerns a Council employee who painted the white lines in the middle of the road. He was hit from behind by a car and injured; he developed fears which were eliminated by behavior therapy (desensitization) to enable him to go back to his job. The success of the therapy was short-lived, however, as he was run down again by another car. He was subjected to desensitization a second time, again successfully, but was finally run over again a third time! We considered it unethical to continue the treatment, as clearly his fears were only part conditioned and irrational, but in large measure also learned and rational. This combination is not unusual, and may be particularly applicable to the studies adduced by De Silva and Rachman as supporting their view.

A proper experimental study of the hypothesis that exposure is not necessary for fear/anxiety reduction, in order to contradict the theory here advocated, would have to show that it dealt with conditioned irrational fears, and did not capitalize on the learned and rational parts of those fears. It is not entirely irrational to be afraid of public speaking or snakes, and hence it seems that the examples given by De Silva and Rachman are contaminated, so that the possible fear reduction through learning may only have affected the learned rational part of the total fear. The position taken by De Silva and Rachman is not necessarily false, but the evidence supporting it is rather insufficient.

The same is true of another example given by them, namely a study by Marks, Gelder, and Edwards (1968) using hypnosis. The hypnotic treatment contained no element of exposure to the feared stimuli, but it did contain relaxation and other general fear/anxiety reducing elements, and in line with Zillmann's theory, outlined earlier, this might reduce the level of anxiety below the critical points (Eysenck, 1982).

One further point that remains to be discussed is the degree to which cognitive factors are opposed to a behavioristic theory of the kind here adopted. Such an opposition would certainly exist in relation to the old-fashioned S-R type of theory adopted by Watson, but not in relation to the more modern S-S type theory (Mackintosh, 1984). The relationship is between stimulus and stimulus, and where, as Pavlov emphasized, words can be used as conditioned stimuli as well as conditioned responses, there is no contradiction involved in explaining results such as those of Wilson (1968), who showed that conditioned responses to a given stimulus could be reversed by suitable instruction (see also Bridger & Mandel, 1964). An even earlier example is Miller's (1935) demonstration that the psychological response could be conditioned to a cognitive stimulus. Miller administered electric shocks to subjects when the letter *T* but not the number *4* was read out aloud, and then instructed

subjects to think *T* and *A*, alternately, in a series of trials. Galvanic skin responses occurred when the subject thought *T*, but not when they thought *A*.

To say all this is not to suggest that exposure is indeed necessary for fear/anxiety reduction, even in the limited sense of conditioned irrational fear/anxiety reduction. It is merely maintained that the evidence against this view is not very strong, and does not take into account considerations that may be vital in assessing the relevance of the studies quoted by the critics. Admittedly it will be difficult to conduct experiments that can establish the reduction of such fears without exposure, but this difficulty should not be allowed to permit the suggestion that the deed had already been done. As far as the existing evidence is concerned, it is robustly in favor of the view that exposure is much the most important, and may be a necessary condition for fear/anxiety reduction. This conclusion may not apply to the same extent to learned fears, whether irrational or rational; much research remains to be done to clear up this particular question.

SUMMARY AND CONCLUSIONS

It may be useful to repeat what is, and what is not, asserted in this chapter. It is asserted that learning theory, and particular modern principles of conditioning and extinction, are basic to the acquisition and treatment of neurotic disorders. It is not asserted that in particular cases other factors may not be of considerable importance, and may be used to help or hinder the development of a proper treatment procedure. However, the principles of conditioning are fundamental in any viable theory of neurosis.

It is not asserted that cognitive therapy (e.g., Hoffman, 1984) constitutes a separate, antagonistic framework of theory and practice contrasted with behavior therapy and the underlying theory of conditioning. It is asserted that modern learning theory, as outlined for instance in the chapter by Dickinson in this volume, takes into account cognitive processes and principles, and combines these in a meaningful manner within learning theory. Information processing (Foa & Kozak, 1986) is an essential part of modern learning theory, as so defined, and does not require us to posit a separate cognitive psychology, separate and apart from theories of learning.

It is not asserted that all human behavior can be reduced to principles derived from animal behavior; no such complete reductionism is intended. It is asserted, however, that certain types of behavior, particularly neurotic behaviors, do find a very close analogue in animal behavior, and that hence the study of conditioning and learning in animals is of fundamental importance for an understanding of the processes mediating the acquisition of neurotic behavior in humans, and its extinction. The Rachman and Hodgson studies of obsessions and compulsions (1980) would seem to establish the correctness of this view once and for all.

It is not asserted that Watsonian theories of neurosis and treatment, or the Hull-Spence type of learning theory, should govern our thinking about the relationship between conditioning and neurosis. It is asserted that these writers laid the foundation for a better understanding of the acquisition of neurotic disorders, and

pioneered an understanding of the principles on which any treatment must be based. Our theories may be in error, but as Francis Bacon remarked, "truth comes out of error more readily than out of confusion."

It is not only our general theory of learning and conditioning that has changed dramatically from an S-R to an S-S theory (or better still, to an S-S-R theory!), but also certain specific applications of learning theory to neurosis. Concepts such as that of incubation of fear seem to be essential if we are to map the facts of neurosis onto the theories of conditioning and learning. It will be a long time before this is done in such a way as to satisfy all our theoretical and practical demands, but already the fit is better than to any other existing theory.

How, in fact, shall we evaluate a theory? To quote Mao Tse Tung: "The only standard by which truth can be assessed is in its practical results." It has been noted in many discussions of the effects of psychotherapy that "all have won, and all must have prizes," a conclusion typical of the Alice in Wonderland state of affairs prevailing in psychotherapeutic research. If indeed all different psychotherapeutic methods, as well as placebo methods, work equally well, then clearly none of the specific theories giving rise to these many different methods of psychotherapy can have any specific value; such effects as are seen must be due either to nonspecific factors, such as suggestion, prestige, friendly human interaction, etc., or as suggested in this chapter, to unintended but nevertheless present Pavlovian extinction. The fact that behavior therapy is surely more effective than psychotherapy or placebo treatment in many cases is clear evidence that specific factors are involved here, and hence that there must be some measure of truth to the principles on which behavior therapy is based.

If that be so, we may perhaps here quote Kurt Lewin's famous saying: "Nothing is more practical than a good theory." Progress in behavior therapy depends crucially on improving and updating our theories in the light of ongoing research, using these theories to improve our method of treatment and checking their efficacy against the effects of treatment. Treatment may be regarded as an extension of laboratory research. We cannot test our theories concerning strong emotions very easily in contrived laboratory settings, for ethical and humane reasons. Neurotic fears present us with a testing bed for predictions derived from our theories. This reciprocal process, laboratory helping clinic, and clinic assisting laboratory, is perhaps the most important outcome of the advance of the behavior therapy movement.

REFERENCES

- Allport, D. A. (1975). The state of cognitive psychology: A critical notice of W. G. Chase (Ed.), *Visual Information Processing*. *Quarterly Journal of Experimental Psychology*, 23, 141-152.
- Armario, A., Castellanos, J. M., & Balasch, J. (1984). Effect of crowding on emotional reactivity in male rats. *Neuroendocrinology*, 39, 330-333.
- Baer, D. M., & Teguchi, H. (1985). Generalized imitation from a radical behavior viewpoint. In S. Reiss & R. R. Bootzin (Eds.), *Theoretical issues in behavior therapy* (pp. 179-217). New York: Academic Press.

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior changes. *Psychological Review*, *84*, 191–215.
- Bandura, A. (1978). Perceived self-efficacy. In S. Rachman (Ed.), *Advances in behavior research and therapy*, *1*, 137–139.
- Barnes, B. (1982). *T. S. Kuhn and Social Science*. London: Macmillan.
- Beck, A. T. (1976). *Cognitive theory and the emotional disorders*. New York: International University Press.
- Beckwith, B. E., & Sandman, C. A. (1978). Behavioral influences of the neuropeptides ACTH and MSH: A methodological review. *Neuroscience & Behavioral Reviews*, *2*, 311–338.
- Beckwith, B. E., & Sandman, C. A. (1982). Central nervous system and peripheral effects of ACTH, MSH, and related neuropeptides. *Peptides*, *3*, 411–420.
- Bertolini, A., & Gessa, G. L. (1981). Behavioral effects of ACTH and MSH peptides. *Journal of Endocrinological Investigations*, *4*, 241–251.
- Blanchard, R. J., Kelley, M. J., & Blanchard, D. C. (1974). Defensive reactions and exploratory behavior in rats. *Journal of Comparative & Physiological Psychology*, *87*, 1129–1134.
- Bootzin, R. R. (1985). Affect and cognition in behavior theory. In S. Reiss & R. R. Bootzin (Eds.), *Theoretical issues in behavior therapy* (pp. 35–45). New York: Academic Press.
- Borrell, J., De Kloet, E. R., Versteeg, D. H. G., & Bohus, B. (1983). Inhibitory avoidance deficit following short-term adrenalectomy in the rat: The role of adrenal catecholamines. *Behavioral & Neural Biology*, *39*, 241–258.
- Breger, L., & McGaugh, J. L. (1965). Critique and reformulation of “learning theory” approaches to psychotherapy and neurosis. *Psychological Bulletin*, *63*, 338–358.
- Bridger, W., & Mandel, A. (1964). A comparison of GSR fear responses produced by threat and electric shock. *Journal of Psychiatric Research*, *2*, 31–40.
- Britton, D. R., & Britton, K. T. (1981). A sensitive open-field measure of anxiolytic drug activity. *Pharmacology Biochemistry & Behaviour*, *15*, 577–582.
- Britton, D. R., Koob, G. F., Rivier, J., & Vale, W. (1982). Intraventricular corticotrophin-releasing factor enhances behavior effects of novelty. *Life Sciences*, *31*, 363–367.
- Bykov, K. M. (1957). *The cerebral cortex and the internal organs*. New York: Chemical Publications.
- Charney, D. S., & Redmond, D. E. Jr., (1983). Biological mechanisms in human anxiety. *Neuropharmacology*, *22*, 1531–1536.
- Concannon, J. T., Riccio, D. C., Maloney, R., & McKelvey, J. (1980). ACTH mediation of learned fear: blockade by naloxone and naltrexone. *Physiology & Behavior*, *25*, 977–979.
- Concannon, J. T., Riccio, D. C., & McKelvey, J. (1980). Pavlovian conditioning of fear based upon hormonal mediation of prior aversive experience. *Animal Learning & Behavior*, *8*, 75–80.
- Davey, G. (1983). *Animal models of human behavior*. London: Wiley.
- Davey, G. (1983). An associative view of human classical conditioning. In G. Davey (Ed.), *Animal models of human behavior* (pp. 95–114). London: Wiley.
- de Silva, P., & Rachman, S. (1981). Is exposure a necessary condition for fear-reduction? *Behavior Research and Therapy*, *19*, 227–232.
- de Vito, W. J., & Brush, F. R. (1984). Effect of ACTH and Vasopressin on extinction: Evidence for opiate mediation. *Behavioral Neuroscience*, *98*, 59–71.
- de Wied, D., & Jolles, J. (1982). Neuropeptides derived from pro-opiocortin: Behavioral, psychological and neurochemical effects. *Psychological Reviews*, *62*, 976–1060.
- Eaglen, A. (1978). Learning theory versus paradigms as the basis for behavior therapy. *Journal of Behavioral Therapy and Experimental Psychiatry*, *9*, 215–218.
- Endroczi, E., Kissak, K., Fekete, T., & de Weid, D. (1970). Effects of ACTH on EEG habituation in human subjects. In D. De Weid & J. A. W. M. Weigen (Eds.), *Pituitary, adrenal and the brain*. Amsterdam: Elsevier.
- Erwin, E. (1978). *Behavior therapy: Scientific, philosophical and moral foundations*. New York: Cambridge Press.
- Eysenck, H. J. (1947). *Dimensions of personality*. London: Kegan Paul.
- Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology*, *16*, 319–324.
- Eysenck, H. J. (1959). Learning theory and behavior therapy. *The Journal of Mental Science*, *105*, 61–75.
- Eysenck, H. J. (Ed.). (1960). *Behavior therapy and the neuroses*. Oxford: Pergamon Press.

- Eysenck, H. J. (Ed.). (1964). *Experiments in behavior therapy*. Oxford: Pergamon Press.
- Eysenck, H. J. (1967). Single trial conditioning neurosis and the Napalkov phenomenon. *Behavior Research and Therapy*, 5, 63–65.
- Eysenck, H. J. (1968). A theory of the incubation of anxiety/fear responses. *Behavior Research and Therapy*, 6, 309–321.
- Eysenck, H. J. (1970a). *The structure of human personality*. London: Methuen.
- Eysenck, H. J. (1970b). A dimensional system of psychodiagnostics. In A. R. Mahrer (Ed.), *New approaches to personality classification* (pp. 167–208). New York: Columbia University Press.
- Eysenck, H. J. (1976). The learning theory model of neurosis—A new approach. *Behavior Research and Therapy*, 14, 251–267.
- Eysenck, H. J. (1979). The conditioning model of neurosis. *The Behavioral and Brain Sciences*, 2, 155–199.
- Eysenck, H. J. (1980). A unified theory of psychotherapy, behavior therapy and spontaneous remission. *Zeitschrift für Psychologie*, 188, 43–56.
- Eysenck, H. J. (1982a). Neobehaviorist (S-R) theory. In G. T. Wilson & C. M. Franks (Eds.), *Contemporary Behavior Therapy* (pp. 205–276). New York: Guilford Press.
- Eysenck, H. J. (1982b). Why do conditioned responses show incrementation, while unconditioned response show habituation? *Behavior Psychotherapy*, 10, 217–222.
- Eysenck, H. J. (1983). Classical conditioning and extinction: The general model for the treatment of neurotic disorders. In M. Rosenbaum, C. Franks, & Y. Jaffe (Eds.), *Perspectives on Behavior Therapy in the Eighties* (Vol. 9, pp. 77–98). New York: Springer.
- Eysenck, H. J. (1985). Incubation theory of fear/anxiety. In S. Reiss & R. R. Bootzin (Eds.), *Theoretical issues in behavior therapy* (pp. 83–105). New York: Academic Press.
- Eysenck, H. J. (1987). Psychotherapy to behavior therapy: A Paradigm shift. In D. B. Fishman, F. Rotgers, & C. Franks (Eds.), *Paradigms in behavior therapy: Present and promise*. New York: Springer.
- Eysenck, H. J., & Beech, H. R. (1971). Counter conditioning and related methods. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (pp. 543–611). New York: Wiley.
- Eysenck, H. J., & Broadhurst, P. L. (1964). Experiments with Animals. In H. J. Eysenck (Ed.), *Experiments in motivation* (pp. 285–291). London: Pergamon Press.
- Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences*. New York: Plenum Press.
- Eysenck, H. J., & Kelley, M. J. (1987). The interaction of neurohormones with Pavlovian A and Pavlovian B conditioning in the causation of neurosis, extinction and incubation of anxiety. In G. Davey (Ed.), *Conditioning in humans*. New York: Wiley.
- Eysenck, H. J., & Rachman, S. (1965). *Causes and cures of neurosis*. London: Routledge & Kegan Paul.
- Eysenck, H. J., Wakefield, J. A., & Friedman, A. (1983). *Diagnosis and clinical assessment: The DSM-III. Annual Review of Psychology*, 34, 167–193.
- Eysenck, M. W. (1984). *A handbook of cognitive processes*. Hillsdale NJ: Erlbaum.
- File, S. E., & Vellucci, S. V. (1978). Studies on the role of ACTH and 5-HT in anxiety using an animal model. *Journal of Pharmacy & Pharmacology*, 30, 105–110.
- Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin*, 99, 20–35.
- Franks, C. M. (1984). On conceptual and technical integrity in psychoanalysis and behavior therapy. In H. Arcowitz & S. B. Messer (Eds.), *Psychoanalytic therapy and behavior therapy*. New York: Plenum Press.
- Gold, P. W., Chrousos, G., Kellner, C., Post, R., Roy, A., Augerinos, P., Schuller, H., Oldfield, E., & Loriaux, D. L. (1984). Psychiatric implications of basic and clinical studies with corticotropin-releasing factor. *The American Journal of Psychiatry*, 141, 619–627.
- Goldberg, D., & Huxley, P. (1980). *Mental illness in the community: A pathway to psychiatric care*. London: Tavistock.
- Goldfried, M. R. (1980). Toward the delineation of therapeutic change. *American Psychologist*, 35, 991–999.
- Gossop, M. (1981). *Theories of neurosis*. New York: Springer.
- Grant, D. A. (1964). Classical and operant conditioning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press.
- Gray, J. A. (1982). *The Neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. New York: Oxford University Press.
- Grey, S., Sartory, G., & Rachman, S. (1979). Synchronous and desynchronous changes during fear reduction. *Behavior Research and Therapy*, 17, 137–147.

- Hadley, S. W., & Strupp. (1976). Contemporary views of negative effects with psychotherapy. *Archives of General Psychiatry*, *33*, 1291–1302.
- Hall, S. M. (1979). The abstinence phobia. In N. A. Krasnegor (Ed.), *Behavioral analysis and treatment of substance abuse* (pp. 55–67). NIDA Res. Monograph, 25, Rockville, MD.
- Haroutunian, V., Riccio, D. C. (1977). Effect of arousal conditions during reinstatement treatment upon learned fear in young rats. *Developmental Psychobiology*, *10*, 25–32.
- Haroutunian, V., & Riccio, D. C. (1979). Drug-induced “arousal” and the effectiveness of CS exposure in the reinstatement of memory. *Behavioral and Neural Biology*, *26*, 115–120.
- Hattie, J., Sharples, C., & Rogers, H. (1984). Comparative effectiveness of professional and paraprofessional helpers. *Psychological Bulletin*, *55*, 534–541.
- Hodgson, R., & Rachman, S. (1974). Desynchrony in measures of fear. *Behavior Research and Therapy*, *12*, 319–326.
- Hoffman, N. (Ed.). (1984). *Foundations of cognitive therapy: Theoretical methods and practical applications*. New York: Plenum Press.
- Iverson, L. L. (1984). Amino acids and peptides: Fast and slow chemical signals in the nervous system. *Proceedings of the Royal Society London*, *221*, 245–260.
- Janet, P. (1890). *L'état mental des hystériques*. Paris: Rueff.
- Janet, P. (1903). *Les obsessions et la psychosthime*. Paris: Alcan.
- Jones, M. C. (1924). The elimination of children's fears. *Journal of Experimental Psychology*, *7*, 383–390.
- Jung, C. G. (1923). *Psychological types*. New York: Harcourt Brace.
- Kalin, N. E., Shelton, S. E., Kraemer, G. W., & McKinney, W. T. (1983a). Associated endocrine, physiological and behavioral changes in Rhesus monkeys after intravenous corticotropin-releasing factor administration. *Peptides*, *4*, 211–215.
- Kalin, N. E., Shelton, S. E., Kraemer, G. W., & McKinney, W. T. (1983b). Corticotropin-releasing factor administered intraventricularly to Rhesus monkeys. *Peptides*, *4*, 217–220.
- Kanfer, F. H., & Hagerman, S. M. (1985). Behavior therapy and the information-processing paradigm. In S. Reiss & R. R. Bootzin (Eds.), *Theoretical issues in behavioral therapy* (pp. 3–33). New York: Academic Press.
- Karseras, A. G. (1976). Psychiatric aspects of ophthalmology: Modern perspectives. In J. G. Howells (Ed.), *The psychiatric aspects of surgery*. London: Macmillan.
- Kelley, M. J. (in press). Epinephrine-cue contiguity is required for the modulation of CS excitatory strength. *Physiological Psychology*.
- Kessel, N. (1960). Psychiatric morbidity in a London general practice. *British Journal of Preventive and Social Medicine*, *14*, 16–22.
- Kessel, N., & Shepherd, M. (1962). Neurosis in hospital and general practice. *Journal of Mental Science*, *108*, 159–166.
- Kimble, G. A. (1961). *Hilgard and Marquis' conditioning and learning*. New York: Appleton-Century-Croft.
- Kimmel, H. D. (1975). Conditioning of fear and anxiety. In C. D. Spielberger & I. G. Sarason (Eds.), *Stress and anxiety* (Vol. 1, pp. 214–231). New York: Wiley.
- Krieger, D. T. (1983). Brain peptides: What, where and why? *Science*, *222*, 975–985.
- Kuhn, T. S. (1959). *The Copernican revolution*. New York: Vintage Books.
- Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.
- Kuhn, T. S. (1974). Second thoughts on paradigms. In F. Suppe (Ed.), *The structure of scientific theories* (pp. 459–482). London: University of Illinois Press.
- Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 194–226). Cambridge: University Press.
- Lang, P. (1970). Stimulus control, response control and the desensitization of fear. In D. Lewis (Ed.), *Learning approaches to therapeutic behavior* (pp. 119–132). Chicago, IL: Aldine.
- Latimer, P. R., & Sweet, A. A. (1984). Cognitive versus behavioral procedures in cognitive behavior therapy: A critical review of the evidence. *Journal of Behavior Therapy and Experimental Psychiatry*, *15*, 9–22.
- Le Roith, D., Shiloach, J., & Roth, J. (1982). Is there an earlier phylogenetic precursor that is common to both the nervous system and the endocrine systems? *Peptides*, *3*, 211–215.
- Levey, A. B., & Martin, I. (1983). Cognition, evaluation and conditioning. Rules of sequence and rules of consequence. *Advances in Behavior Research and Therapy*, *4*, 181–195.

- Locke, E. A. (1971). Is "behavior therapy" behavioristic? *Psychological Bulletin*, *76*, 318–327.
- London, P. (1972). The end of ideology in behavior modification. *American Psychologist*, *27*, 913–920.
- Mackintosh, N. J. (1974). *The psychology of animal learning*. London: Academic Press.
- Mackintosh, N. J. (1984). *Conditioning and associative learning*. Oxford: Clarendon Press.
- Mackenzie, B. D. (1977). *Behaviorism and the limits of scientific method*. London: Routledge & Kegan Paul.
- Mahoney, M. J. (1974). *Cognition and behavior modification*. Cambridge, MA: Ballinger.
- Markey, K. A., & Sze, P. Y. (1984). Influence of ACTH on tyrosine hydroxylase activity in the locus coeruleus of mouse brain. *Neuroendocrinology*, *38*, 269–275.
- Marks, I. M., Gelder, M. G., & Edwards, G. (1968). Hypnosis and desensitization for phobias: A controlled prospective trial. *British Journal of Psychiatry*, *114*, 1263–1274.
- Martin, I., & Levey, A. B. (1985). Conditioning, evaluations and cognitions: An axis of integration. *Behavior Research and Therapy*, *23*, 167–175.
- Marzillier, J. S. (1980). Cognitive therapy and behavioral practice. *Behavior Research and Therapy*, *18*, 249–258.
- Matt, G. E., & Wittmann, W. W. (1985). Zum Status der kontrollierten deutschsprachigen Psychotherapie—Effektforschung aus dem Blickwinkel einer Meta-Analyse. *Zeitschrift für Klinische Psychologie*, *14*, 293–312.
- McGaugh, J. L. (1983). Hormonal influences on memory. *Annual Review of Psychology*, *34*, 297–323.
- Meichenbaum, D. (1977). *Cognitive behavior modification: An integrative approach*. New York: Plenum Press.
- Meichenbaum, D. H., Gilmore, J. B., & Fedoravicius, A. (1971). Group insight versus group desensitization in treating speech anxiety. *Journal of Consulting and Clinical Psychology*, *36*, 410–421.
- Mendelsohn, K. (1966). *The quest for absolute zero*. London: Weidenfeld & Nicolson.
- Messer, S. B., & Winokur, M. (1980). Some limits of the integration of psychoanalytic and behavior therapy. *American Psychologist*, *35*, 818–827.
- Miller, N. E. (1935). The influence of past experience upon the transfer of subsequent training. Unpublished doctoral dissertation, Yale University, New Haven, CT.
- Morley, S. (1977). The incubation of avoidance behavior: Strain differences in susceptibility. *Behavior Research and Therapy*, *15*, 365–367.
- Mowrer, O. H. (1948). Learning theory and the neurotic paradox. *American Journal of Orthopsychiatry*, *18*, 571–610.
- Mowrer, O. H. (1950). *Learning theory and personality dynamics*. New York: Arnold Press.
- Murray, E. J., & Foote, F. (1979). The origins of fear of snakes. *Behavior Research and Therapy*, *17*, 489–493.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we know: Verbal reports on mental processes. *Psychological Review*, *84*, 231–259.
- Office of Population and Censuses Surveys. (1974). *Morbidity Statistics from general practice. Second national study, 1970–71*. London: HMSO.
- Öhman, A., Dimberg, V., & Öst, L.-G. (1985). Animal and social phobias: Biological constraints on learned fear responses. In S. Reiss & R. R. Bootzin (Eds.), *Theoretical issues in behavior therapy* (pp. 123–175). New York: Academic Press.
- Öst, L.-G., & Hugdahl, K. (1981). Acquisition of phobia and anxiety response patterns in clinical patients. *Behavior Research and Therapy*, *19*, 439–477.
- Pagano, R. R., & Lovely, R. H. (1972). Diurnal cycle and ACTH facilitation of shuttlebox avoidance. *Physiology & Behavior*, *8*, 721–723.
- Pavlov, I. P. (1927). *Conditioned reflexes*. London: Oxford Press.
- Platanov, K. (1959). *The word as a physiological and therapeutic factor*. Moscow: Foreign Languages Publishing House.
- Prioleau, L., Murdock, M., & Brody, N. (1983). An analysis of psychotherapy versus placebo studies. *The Behavioral and Brain Sciences*, *6*, 275–285.
- Rachman, S. (1977). The conditioning theory of fear-acquisition: A Critical examination. *Behavior Research and Therapy*, *15*, 375–387.
- Rachman, S. (1978). *Fear and courage*. San Francisco: Freeman.
- Rachman, S. (1981). The primacy of affect: Some theoretical implications. *Behavior Research and Therapy*, *19*, 279–290.
- Rachman, S., & Hodgson, R. (1974). Synchrony and desynchrony in fear and avoidance. *Behavior Research and Therapy*, *12*, 311–318.

- Rachman, S., & Hodgson, R. (1980). *Obsessions and compulsions*. Englewood Cliffs, NJ: Prentice-Hall.
- Rachman, S., & Wilson, G. T. (1980). *The effects of psychological therapy*. London: Pergamon Press.
- Redmond, D. E., Jr. (1981). Clonidine and the primate locus coeruleus: Evidence suggesting anxiolytic and anti-withdrawal effects. In H. Lal & S. Fielding (Eds.), *Psychopharmacology of clonidine* (pp. 147–153). New York: Alan Liss.
- Redmond, D. E., Jr., & Huang, Y. H. (1979). New evidence for a locus coeruleus-norepinephrine connection with anxiety. *Life Sciences*, *25*, 2149–2162.
- Redmond, D. E., Jr., & Krystal, J. H. (1984). Multiple mechanisms of withdrawal from opioid drugs. *Annual Review of Neuroscience*, *7*, 443–478.
- Reiman, E. M., Raichle, M. E., Butler, F. K., Herscovitch, P., & Robins, E. (1984). A focal brain abnormality in panic disorder, a severe form of anxiety. *Nature*, *310*, 683–685.
- Reiss, S. (1980). Pavlovian conditioning and human fear: An expectancy model. *Behavior Therapy*, *11*, 380–396.
- Rescorla, R. A. (1972). Informational variables in Pavlovian conditioning. In G. M. Bower (Ed.), *The psychology of learning and motivation*. New York: Academic Press.
- Righter, H., Elbertse, R., & van Riezen, H. (1975). Time-dependent anti-amnesic effect of ACTH 4-10 and desglycinamide-lysine vasopressin. *Progress in Brain Research*, *42*, 163–171.
- Rimm, D. C., Janda, L. H., Lancaster, D. W., Nake, M., & Dittmar, K. (1977). An exploratory investigation of the origins and maintenance of phobias. *Behavior Research and Therapy*, *15*, 231–238.
- Rimm, D. C., & Masters, J. C. (1979). *Behavior therapy: Techniques and empirical findings* (2nd ed.). New York: Academic Press.
- Searles, J. S. (1985). A methodological and empirical critique of psychotherapy outcome meta-analysis. *Behavior Research and Therapy*, *23*, 453–463.
- Shapiro, D., & Shapiro, D. (1982). Meta-analysis of comparative therapy outcome studies: A replication and refinement. *Psychological Bulletin*, *92*, 581–604.
- Simms, A. (1983). *Neurosis in society*. Basingstoke: Macmillan.
- Simms, A. (1985). Neurosis 1784–1984: Two centuries of obfuscation. *British Journal of Clinical and Social Psychiatry*, *3*, 24–26.
- Simms, A., & Sandman, P. (1975). The severity of symptoms of psychiatric outpatients: Use of the General Health Questionnaire in hospital and general practice patients. *Psychological Medicine*, *5*, 62–66.
- Smith, M. L., Glass, G. V., & Miller, T. I. (1980). *The benefits of psychotherapy*. Baltimore, MD: John Hopkins University Press.
- Staats, A. W. (1964). *Human learning*. New York: Holt, Rinehart & Winston.
- Staats, A. W. (1968). *Learning, language and cognition*. New York: Holt, Rinehart & Winston.
- Strupp, H. H., Hadley, S. W., & Gomes-Schwartz, B. (1977). *Psychotherapy for better or worse: The problem of negative effects*. New York: Aronson.
- Sutherland, S. (1976). *Breakdown*. London: Weidenfeld.
- Sutherland, S., & Mackintosh, N. J. (1971). *Mechanisms of animal discrimination Learning*. New York: Academic Press.
- Taylor, J. G. (1962). *The behavioral basis of personality*. New Haven, CT: Yale University Press.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*, 189–208.
- Tyron, W. W. (1978). An operant explanation of Mowrer's neurotic paradox. *Behaviorism*, *6*, 203–211.
- Ullmann, L. P. (1981). Cognition: Help or hindrance? *Journal of Behavior Therapy and Experimental Psychiatry*, *12*, 19–24.
- Urban, I. J. A. (1984). Electrophysiological effects of peptides derived from pro-opiomelanocortin. *Pharmacology & Therapeutics*, *24*, 57–90.
- Urban, I. J. A., & de Weid, D. (1978). Neuropeptides: Effects on paradoxical sleep and theta rhythm in rats. *Pharmacology Biochemistry & Behavior*, *8*, 51–59.
- van Wimersma Greidanus, Tj. B., Bohus, B., Kovacs, G. L., Versteeg, D. H. G., Burbach, J. P. H., & de Wied, D. (1983). Sites of behavioral and neurochemical action of ACTH-like peptides and neurohypophysial hormones. *Neuroscience & Biobehavior Reviews*, *7*, 453–463.
- van Wimersma Greidanus, Tj. B., Rees, L. H., Scott, A. D., Lowry, P. J., & de Wied, D. (1977). ACTH release during passive avoidance behavior. *Brain Research Bulletin*, *2*, 101–104.
- Wachtel, P. L. (1977). *Psychoanalysis and behavior therapy*. New York: Basic Books.

- Wagner, A. P., & Rescorla, R. A. (1972). Inhibition in Pavlovian conditioning: Application of a theory. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and Learning*. (pp. 224–238). New York: Academic Press.
- Watson, J. B., & Rayner, R. (1920) Conditioned emotional reactions. *Journal of Experimental Psychology*, *3*, 1–14.
- Weinberger, N. M., Gold, P. E., & Sternberg, D. B. (1984). Epinephrine enables Pavlovian fear conditioning under anesthesia. *Science*, *223*, 605–607.
- Weiss, K. S., Nelson, R. O., & Odom, J. V. (1975). The relative contribution of reattribution and verbal extinction to the effectiveness of cognitive restructuring. *Behavior Therapy*, *6*, 459–476.
- Weissberg, M. (1977). A comparison of direct and vicarious treatment of speech anxiety: Desensitization, desensitization with coping imagery, and cognitive modification. *Behavior Therapy*, *8*, 606–620.
- Wilson, G. D. (1968). Reversal of differential GSR conditioning by instructions. *Journal of Experimental Psychology*, *76*, 491–493.
- Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.
- Wolpe, J. (1973). *The practice of behavior therapy* (2nd ed.). New York: Pergamon Press.
- Wolpe, J. (1976a). Behavior therapy and its malcontents—1. Denial of its bases and psychodynamic fusionism. *Journal of Behavioral Therapy & Experimental Psychiatry*, *7*, 1–5.
- Wolpe, J. (1976b). Behavior therapy and its malcontents. 11. Multimodel eclecticism, cognitive exclusionism and exposure empiricism. *Journal of Behavior Therapy & Experimental Psychiatry*, *7*, 109–116.
- Wolpe, J. (1978). Cognition and causation in human behavior and its therapy. *American Psychologist*, *33*, 437–446.
- Wolpe, J. (1981). Behavior therapy versus psychoanalysis. *American Psychologist*, *36*, 159–164.
- Woods, P. J. (1974). A taxonomy of instrumental conditioning. *American Psychologist*, *29*, 585–597.
- Yates, A. J. (1983). Behavior therapy and psychodynamic psychotherapy: Basic conflict or reconciliation and integration? *British Journal of Clinical Psychology*, *22*, 107–125.
- York, C. (1966). *If hopes were dupes*. London: Hutchinson.
- Zajonc, R. B. (1980). Feeling and thinking. *American Psychologist*, *235*, 151–175.
- Zajonc, R. B. (1984). On the primacy of affect. *American Psychologist*, *39*, 17–123.
- Zener, K. (1937). The significance of behaviour accompanying conditioned salivary secretion for theories of the conditioned response. *American Journal of Psychology*, *50*, 384–403.
- Zigmond, A., & Simms, A. (1983). The effect of the use of the International Statistical Classification of Diseases, 9th Revision, upon hospital diagnoses. *British Journal of Psychiatry*, *142*, 409–413.
- Zillmann, D. (1978). Attribution and misattribution of excitatory reaction. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), *New directions in the attribution research (tools)*. Hillsdale, NJ: Erlbaum.
- Zillmann, D. (1979). *Hostility and aggression*. Hillsdale, NJ: Lawrence Erlbaum.
- Zillmann, D. (1984). *Connections between sex and aggression*, (pp. 184–193). Hillsdale, NJ: Erlbaum.
- Zuriff, G. E. (1985). *Behaviorism: A conceptual reconstruction*. New York: Columbia University Press.

CHAPTER 2

Behavior Modification

Angela Schorr

DEFINITION OF TERMS AND FUNDAMENTAL PRINCIPLES

This chapter deals with the application of the methods of operant learning theory, here called behavior modification, to clinical and social problems. As behavior therapy developed, there were numerous attempts to distinguish between the terms *behavior therapy* and *behavior modification* (e.g., Franzini & Tilker, 1972; Keehn & Webster, 1969; Pomerleau, 1979). Standard works on behavior modification still contain traces of such subtle distinctions in meaning although the case for using the terms synonymously has been put authoritatively (cf. Kazdin, 1978, 1984; Krasner, 1971; Mahoney, Kazdin, & Lesswing, 1974). In their introductory textbook, *Behavior Modification in the Human Services*, Sundel and Sundel (1982, p. 280) define behavior modification as the “application of principles and techniques derived from the experimental analysis of behavior to a wide range of human problems.” Behavior modification is based on the methods of applied behavior analysis, the principles of operant conditioning, and the conceptual framework of social learning theory. The authors define behavior therapy, on the other hand, as follows:

[It is] generally used as synonymous with the term behavior modification. The term behavior therapy connotes the provision of behavior modification services to individuals in a client-therapist setting. Historically, behavior therapy referred to the treatment methods based primarily on classical conditioning. (Sundel & Sundel, 1982, p. 280)

This definition is one indication of the persistent individualism characteristic of Skinner’s behavior therapy disciples. In his 1984 presidential address to the Association for the Advancement of Behavior Therapy, Ross for this reason chose the term “our *cousins* in applied behavior analysis” to refer to this school of behavior therapy (Ross, 1985, p. 198; my emphasis). Unlike any other behavior therapy school, operant behavior therapists in the United States can point to two traditional groupings, the Association for the Advancement of Behavior Therapy (AABT) and Division 25 of the American Psychological Association, the Division of the Experimental

Analysis of Behavior, which has produced numerous pioneers in the field of applied behavior analysis whose work is still regarded as standard.

Applied behavior analysts, or behavior modifiers, have always endeavored to stress the scientific nature of their work. Lazarus (1971), for instance, recalls that they soon abandoned the term *therapy*, calling themselves instead “behavior modifiers.” The latter term is comprehensive enough to cover the extensive application of operant techniques to education, treatment, and rehabilitation. The historical development of behavior modification can be described as a “relatively clear progression from experimentation to application” (Kazdin, 1978), a continuous progression from basic research (i.e., experimental analysis of behavior) to applied research, as shown in the work of men like Ferster, Azrin, and Lindsley. The debt owed to basic research is acknowledged in the term *applied behavior analysis*, which came into use when the *Journal of Applied Behavior Analysis* was founded. The term implies both scientific standing and continuity.

Skinner’s theory of operant conditioning is fundamental doctrine for both experimental and applied behavior analysts. The basic tenets of this theory can be outlined briefly as follows: The object of psychology is, according to Skinner, the study of behavior. Research should, as far as possible, be descriptive without being theoretical (Skinner, 1938, 1953, 1977). In *The Behavior of Organisms* (1938) Skinner criticized the use of intervening variables and hypothetical constructs in psychological research. Moreover, with reference to Hull’s hypothetico-deductive model, which was standard at the time, Skinner declared that it was not the goal of research to combine as many theoretical approaches as possible in one comprehensive theory. On the contrary, psychological research should concentrate on the object of observation. He pointedly belittled theoretical approaches as “any explanation of observed fact which appeals to events taking place somewhere else, at some other level of observation described in different terms, and measured, if at all, in different dimensions” (Skinner, 1950, p. 193).

Granted that Skinner’s attitude to theory, as it is still presented in nearly all standard works on operant learning theory and behavior modification, was seminal for the development of operant research. Today, however, it may be advisable to subject it to historical interpretation. Skinner’s rejection of theoretical models reflected to a degree his disappointment in the models of his time, particularly that of Hull, the former leading exponent of scientific psychology. Hull tried to integrate utterly different approaches, in particular learning theories and psychoanalysis in his model, which became almost incomprehensibly formalistic (Schorr, 1984a). Aware of the deficiencies inherent in this approach, Skinner (1938) called for a consistently inductive instead of a hypothetico-deductive approach. Although he realized that conclusions could never be reached empirically without some theoretical foundation, he aimed at a methodological approach employing only those concepts needed for presenting a large number of experimental facts (Ferster & Skinner, 1957; Zuriff, 1985).

Skinner did not distinguish between respondent behavior (Type S conditioning) and operant behavior (Type B conditioning), a type of spontaneous reaction (Skinner, 1935, 1937, 1938), until the late 1930s. Since then, he has concentrated on operant behavior, specifically the relationship between operant behavior and its consequences

as well as the possibilities of manipulating it. Skinner called *behavior frequency*, or probability, recorded cumulatively as a function of time, “a natural datum in a science of behavior” (Skinner, 1970, p. 3). He used the experimental analysis of behavior to study the fundamental principles of reinforcement, punishment, extinction, discrimination, stimulus control, and so forth (Skinner, 1970; for an explanation, cf. Kazdin, 1978, 1984). Intrasubject-replication designs enabled Skinner to draw conclusions from the observation of very few organisms or even from one. Experiments were conducted to compensate for interindividual variability by improving and carefully controlling experimental conditions instead of manipulating group size. Although operant behavior researchers are now acknowledged even by experimental psychologists as “bringers of baselines” (Marr, 1984), their experimental single-case designs and rejection of statistics were opposed by conservative experimental psychologists in the 1950s and 1960s (Kazdin, 1978; Krantz, 1971).

By the early 1950s Skinner was investigating neurotic, psychotic, and disturbed behavior. By then, research for *Schedules of Reinforcement* (Ferster & Skinner, 1957) had been almost completed. The authors felt they had a theory with limitless potential. The objective of the second phase was to demonstrate the validity and broad applicability of their laboratory procedures. Skinner stressed the role of environment in the etiology of mental disorder. He was aware that heredity as well as organic factors like endocrine malfunction and pharmacological influences were etiologically significant. Nevertheless, he thought that

modes of behavior characteristic of mental disease may be simply the result of a history of reinforcement, an unusual condition of deprivation or satiation, or an emotionally exciting circumstance. Except for the fact that they are troublesome or dangerous, they may not be distinguishable from the rest of the behavior of the individual. (1961, p. 198)

Engaged as he was in operant behavior research, that is, the study of behavior modification and maintenance through environmental factors, Skinner did not expand his conception of the etiology of mental disease. His conclusions, which he regarded as tentative, have nevertheless remained valid for behavior modification. Kazdin conjectures that

individuals in a therapeutic environment have not responded to the somewhat irregular contingencies operative in the ‘real world’. If reinforcement were frequent enough for these individuals or delivered in a systematic fashion in ordinary social interaction, the use of behavior modification techniques might not be required. (Kazdin, 1980, pp. 67–68)

From this basis he deduces reinforcement maximization (praise, accomplishment, esteem from others, self-esteem, social interaction) and punishment minimization (stigma, social censure, self-depreciation, repeated failure) as the principal aim of therapy (cf. Kazdin, 1984)

Both experimental behavior analysis and applied behavior analysis derive from a common intellectual and structural frame of reference. Several noticeable differences between the two groups, however, led applied behavior analysts in the late 1960s to new departures. In 1958 Skinner and his disciples had founded the *Journal of Experimental Analysis of Behavior* (*JEAB*). Because of their methodological peculiarities, operant behaviorists had had little chance to publish in the traditional journals of experimental psychology. Ten years later, there was need of yet another journal, the

Journal of Applied Behavior Analysis (JABA). The editors of *JEAB* could not keep up with the rapid increase in the number of behavior modification studies. Moreover, they preferred to maintain the character of the journal as a forum for basic experimental research (Kazdin, 1978). It became apparent, furthermore, that operant behavior therapists were unable to fulfill the methodological requirements made by colleagues working mainly with animal experiments.

The founding of the new journal represented a significant development in the field of behavior modification. The departure from basic research which had already occurred in empirical studies was acknowledged. Applied research and basic research in the operant paradigm differ not only in the degree of methodological precision reached and the use of human instead of animal experimental subjects. They are also essentially different in their aims and areas of research. The shift of focus in research has been described systematically in a seminal article by Baer, Wolf, and Risley (1968). The authors define applied behavior analysis as follows:

the process of applying sometimes tentative principles of behavior to the improvement of specific behaviors, and simultaneously evaluating whether or not any changes noted are indeed attributable to the process of application. (1968, p. 91)

By emphasizing the primacy of research in the field, the authors establish in this essay new research goals. Because it now aims to *improve* specific behavior patterns, behavior modification research no longer deals with all types of behavior and their variables, but only with variables that will improve behavior effectively. Despite the difficulties involved, behavior modification research must focus on behavior patterns of social significance. Applied behavior studies are often conducted in social settings instead of in the laboratory. The authors concede that it has thus become more difficult to establish experimental controls, which in turn has led to an insoluble "problem of judgement" (Baer *et al.*, 1968).

Adherence to a common tradition and a high degree of interchangeability between the two groups (cf. Krantz, 1971) are the definitive characteristics of operant behaviorists. Their approach is distinctive in the field of behavior therapy in that it follows clearly defined theoretical and methodological guidelines. Kazdin and Wilson (1978) describe the task of applied behavior analysis thus: Within the conceptual framework of operant learning theory, applied behavior analysis seeks to examine behavior disorders as a function of their consequences. Therapy usually aims to modify overt behavior but does not seek to alter cognitive processes and private events. The effects of treatment are evaluated by means of single-case experimental designs. One of the questions dealt with in the following section is whether the above description is still valid.

PRACTICAL APPLICATION: A ONE-WAY STREET? SCIENCE AND TECHNOLOGY IN APPLIED BEHAVIOR ANALYSIS

Operant behavior modification techniques are being applied to an increasingly wider range of societal contingencies. Their use in everyday behavior settings is also growing. A glance through the most recent issues of *JABA* (1980–1985) confirms the

impression of an increase in publications on measures for traffic safety, for instance, or on dealing with environmental issues such as energy conservation, recycling, and pollution problems. The number of publications on behavior modification in the field of clinical medicine and on health care issues has increased as well. In the early 1980s the catchword "environmental design" stood for further application of operant behavior modification techniques approached from a new angle. This approach combines a conceptual framework derived from applied behavior analysis and social learning theory with elements of social psychology such as open education, architectural planning, and social engineering (Krasner, 1980; see also Nietzel, Winett, MacDonald, & Davidson, 1977).

At the same time, it has become apparent that applied behavior analysts are becoming less willing to adhere to techniques of behavior change prescribed by operant learning theory. Like the proponents of other schools of behavior therapy, they have become less formalistic in their approach. Despite the primacy of operant learning procedures, eclectic therapy approaches are as common in the field as elsewhere, when they seem appropriate and their effectiveness can be evaluated. Azrin demonstrated in 1977 that therapy could not be based on the concept of reinforcement alone. Referring to what he had experienced while researching and developing token economy programs and other projects, he claimed:

For all of the treatments, changes in the initial reinforcement conception were required, and although the additional procedures were often derived in turn from a reinforcement framework, their necessity was not predicted by the model. The final treatment in every case required improvisations, detours, and innovations because of problems unanticipated by the reinforcement analysis. (p. 143)

By the late 1970s and early 1980s, operant behavior therapists were being required not merely to prove the purely short-term effectiveness of their procedures as compared to other forms of therapy. In addition, they were compelled to develop realistic long-term courses of treatment for helping patients to overcome problems on an individual basis. Operant behavior analysis has fulfilled these requirements and is developing technical aids increasingly geared to use in therapy itself. Despite numerous carefully corroborated new developments, a growing host of critics is warning against a "drift into technology" and expressing doubts about the scientific probity of the field (cf. Deitz, 1978; Hayes, Rincover, & Solnick, 1980; Ross, 1985). In 1978 Deitz expressed concern that "the field seems to be shifting from *applied* behavior analysis to *applying* behavior analysis" (p. 807). After a thorough study of the first ten issues of the *JABA*, Hayes *et al.* (1980) concluded that methodological questions were losing ground to purely technical aspects. According to Baer *et al.* (1968):

The field of applied behavior analysis will probably advance best if the published descriptions of its procedures not only are precisely technological, but also strive for relevance to principle. This can have the effect of making a body of technology into a discipline rather than a collection of tricks. Collections of tricks historically have been difficult to expand systematically and, when they are extensive, difficult to learn and teach. (p. 96)

Ross cautioned the AABT in 1984 that because there had been no serious groundwork in behavior therapy for a long time, there were indications of its indeed becoming a "collection of tricks":

We should not be in the position of having developed a technique for which we must now seek a theoretical rationale. (Ross, 1985, p. 203)

In their preface to *Failures of Behavior Therapy*, a critical, up-to-date assessment of the field intended to stimulate new hypotheses and to explore new ideas, Foa and Emmelkamp (1983) agree that “although behavioral therapy had developed rapidly in the last decade, few new ideas had been advanced recently.” Deitz (1978) in turn is concerned that, because of the emphasis placed on the practical applications of applied behavior analysis,

the investigative, analytic aspects of applied behavior analysis are subordinated to its useful, applicable aspects. This is a change from investigating independent variables to improving (curing) dependent variables. Above all, the change is one that moves the field from primarily a research (scientific) endeavor to primarily one of implementation (technological). (p. 806–807)

In diagnosing a basic shift in the objectives of the field as a whole, Deitz is referring to statements made by leading applied behavior analysts like Azrin, who in 1977 described his own work as “outcome-oriented” and “consumer-directed” (p. 148). Deitz predicts serious consequences for scientific progress:

Scientific research will be replaced by technological demonstrations. New information can be gained from seeing effects on different dependent variables or by studying implementation, but this new information is of a noticeably different type and may not be as useful to a science of behavior. (1978, p. 807)

The critical state of cognitive behavior modification noted by many specialists may serve as a warning of what can happen when development is too dependent on technological advances (cf. Eysenck, in this volume; Ledwidge, 1978, 1979; Mahoney & Kazdin, 1979; Ross, 1985). Mahoney and Arnkoff (1978) regard the lack of formal theoretical models as the essential reason for the inadequacy of cognitive learning approaches:

It is difficult to evaluate a perspective critically when its specific hypotheses and predictions are not clearly delineated. Although flexibility and informality may be the adaptive features in the early gestation of a model, its later growth and refinement require the more focused scrutiny that can only be provided in the context of a formal paradigm. (p. 712)

Ross adds:

I believe that cognitive therapy was not necessarily a step in the wrong direction, as Ledwidge (1978) once argued, but surely it was a precipitately premature step because it was taken before the necessary theoretical foundations were available. (1985, p. 199)

A lack of theoretical progress and a strongly technological research bias are interdependent. However, a specific explanation for the stagnation in conceptual development shown by the field of applied behavior analysis is discernible in addition to the general trend remarked in behavior therapy by Ross (1985). If one looks closely at what operant behaviorists say, one can fault their arguments. Let us first consider statements on the possible correlation of cognitive factors and private events with behavior. Skinner begins his article “Why I Am Not a Cognitive Psychologist” by stating dogmatically: “The variables of which human behavior is a function lie in the environment” (Skinner, 1977, p. 1). Skinner’s radical behaviorism does not ascribe a mediating function to the psyche, which he sees as a product of the interaction of

the individual with environmental contingencies (Skinner, 1945, 1957, 1974, 1977, 1986). Or, as Marr (1984) puts it:

It was the world outside the skin that gave us a world inside the skin. Because private events (e.g. covert behavior) were placed no more than on an equal footing with public behavior, they could not properly be considered *fundamental causes* of public behavior. (p. 356)

Because private events are regarded as equivalent to empirically observable public behavior, methodological requirements render the question of internal sources of behavioral variability irrelevant. An example of this is the work of Biglan and Kass (1977), who take cognitive as well as private events into consideration by classifying "any organismic event" as behavior. They proceed to limit the potential significance of their analysis drastically by insisting that:

With respect to private events, it is also necessary for the investigator to provide corroborative evidence that the event occurred at all. This requirement is not different in kind from the requirement for accuracy in the observation of accessible events. (p. 4)

The authors thus aim to maintain research procedures in theory and methodology guaranteeing customary standards of accuracy in regard to subject matter and influencing factors. Deitz gives a classic demonstration of the way in which this objective can become an end in itself. He remarks on the efforts of cognitive behaviorists in respect to self-control, self-reinforcement, perception, and cognition:

The behavioristic philosophy of science specifies the assumption that the sources of variability are external. It states that behavior is a function of contemporary and historical interactions with the environment (Reynolds, 1975). I think those external sources of variability have not been explored in sufficient detail to claim that some form of intrinsic variability exists. (Deitz, 1978, p. 810)

His introduction to these remarks was to insist: "It is not my point, however, to dismiss or ignore internal variables."

Dogmatic claims supported by sophisticated operant behavior research strategies are not a very convincing basis for exploring the influences of cognition on behavior with an open mind. This attitude not only prevents further exploration of the actual contribution made by internal factors to behavioral variability. It also contributes substantially to the general trend away from theory toward technology. This becomes particularly obvious in Azrin's article "Reevaluation of Basic Assumptions," where he states:

One basic assumption of the learning therapies is the primacy of overt behavior rather than insight or subjective events. Yet in all of the treatments I found it necessary to include procedures for questioning the trainees or patients to help discover their individual reinforcers. Subjective measures were obtained as adjuncts to the behavioral measures, the strategy being to include such subjective dimensions if the clinical reality required them in spite of the a priori behaviorist proscription against them. (1977, p. 145)

Conceptual inadequacy in applied behavior analysis does not result in a failure to act. On the contrary, it leads to purely technological advances, to an acceptance of treatment procedures which can be evaluated only in terms of effectiveness. These technological developments can be reintegrated in basic and applied research only to a limited extent. A further obstacle to enlarging the conceptual foundations of applied behavior analysis is the unquestioned theoretical bias of operant research methodology. This prevents any modification of the theoretical framework and ensures

the continued inadequacy of conceptual approaches. As Shimp (1984) put it: "Virtually no one explicitly espouses S-R associationism or any other kind of associationism" (p. 414). Yet adherence to associationism is a basic tenet of classical laboratory procedures in the experimental analysis of behavior:

Historically the rôle of theory in the experimental analysis of behavior has been represented substantively by an associative theory implicit in experimental methods and formally by an inductive, algebraic approach. (Shimp, 1984, p. 414)

The way in which the question of individual variability in the operant paradigm is dealt with shows clearly how mixing theory with methodology can lead to the emergence of "anomalies" (cf. Kuhn, 1976, 1977). By rigorously controlling experimental conditions and restricting themselves to the use of animals as experimental subjects, operant behaviorists have been able to minimize the significance of individual variability. Harzem (1984) points out that this used to pose problems in experiments with human subjects:

Often, individual differences have appeared to be the dominant feature of the data obtained in human experiments, resisting the "smoothing" effect of the powerful techniques of behavior analysis. (p. 388)

In a seminal article commemorating the twenty-fifth anniversary of the foundation of the *JEAB*, Harzem suggests that operant behavior analysts should study individual differences and personality:

Indeed, after decades of behavior analysis it is clear that individual differences in behavior cannot be entirely eliminated; they pervade all of the research literature to a greater or lesser extent. (1984, p. 387)

Yet another obstacle to broadening the conceptual framework of applied behavior analysis that should be mentioned here is the dogmatic attitude of operant paradigm researchers to theoretical approaches in general. Again, a second glance at the experimental analysis of behavior can throw light on the current situation in the field. Marr (1984) is critical of the fact that the work of many operant behaviorists has remained on a purely methodological level simply because they are afraid of getting involved with theory. In his article "Cognition, Behavior and the Experimental Analysis of Behavior" (1984), Shimp, too, severely criticized the skeptical attitude of operant behaviorists to theory as well as Skinner's postulate of a purely inductive procedure. Finally, the growing aversion to "explanatory fictions" forced the editors of the *JEAB* in 1980 to admonish contributing authors in an editorial against merely collecting data on behavior determinants instead of dealing with subjects of systematic significance (Nevin, 1980). Conservative applied behaviorists like Deitz (1978) and Ferster (1978), on the other hand, continue to call for a rejection of hypothetico-deductive procedures. Ferster, a pioneer in applied behavior analysis, deplores the trend in basic and applied research to mentalistic imprecision instead of precise description as shown in the *Handbook of Operant Behavior* (Honig & Staddon, 1977). He interprets this development as "a very marked return to the theoretical style of pre-Skinner psychology" (Ferster, 1978, p. 348).

Trends in operant research are being discussed more openly and aggressively in basic research than by applied behavior analysts. In the series of articles on current

and projected developments in experimental behavior analysis commemorating the twenty-fifth anniversary of the *JEAB* referred to above, Neuringer (1984) contends:

In general the field has been narrow, and the basic research questions studied for the past 50 years have been of little interest to other than operant conditioners. (p. 399)

Marr complains, “The experimental analysis of behavior has lagged far behind mainstream psychology, particularly cognitive psychology, in the study of complex behavior” (1984, p. 353). Looking at this state of affairs through Skinner’s eyes has misled quite a few analysts into assuming that complex behavior had been thoroughly investigated. Marr adds:

We have too long and too extensively relied on our founder, B. F. Skinner, as an authority to explain and even to establish the significance of a behavior issue. (p. 361)

Applied behavior analysts, on the other hand, tend to adopt a defensive strategy. Indeed, they seem to have little use for theoretical approaches. Kazdin defines the field of behavior therapy as follows: “Behavior modification is best defined by a rationale and a methodology and not by a specified theory or set of principles” (1984). By 1970 Risley was suggesting that

behavior modification should be viewed as the experimental investigation of socially significant behaviors and not as a simple application of already known principles (i.e., theory) to the therapy situation.

Like Deitz (1978) and Ross (1985), I doubt whether progress in behavior therapy, especially in the area of behavior modification, can be achieved in the long run simply by making it more effective and concentrating on technologically oriented research, however useful or necessary such an approach may seem at present. “It is time to stop standing still!” declared Ross to the members of the AABT in 1984. The various theoretical approaches must be systematically analyzed and integrated (see Eysenck, in this volume). Attempting to achieve this aim primarily on a pragmatic basis, that is, through the application and evaluation of each therapeutic method, could paralyze scientific development. Theoretical advances in behavior therapy can be achieved only if dogmatic claims are rejected out of hand. Moreover, applied behavior analysts would be well advised to cooperate with experimental behavior analysts—their natural allies, by virtue of a common tradition—in dealing with theoretical questions. Neuringer sees advantages for both groups in cooperation: “Applied tests would help the behavior analytic field avoid self-contained, and possibly barren, areas” (1984, p. 400). Fundamental questions that applied research hasn’t had time to pursue could thus be systematically investigated.

THE TOKEN ECONOMY: AN ACHIEVEMENT-ORIENTED SOCIETY MADE IN THE BEHAVIORIST LABORATORY

Token economies like those first developed by Azrin and Ayllon in a ward of the state psychiatric clinic in Anna, Illinois, may well be the most exemplary offshoot of applied behavior analysis. The existing environment is transformed into a laboratory.

Restructuring a given environment ensures that the principles of operant conditioning can regulate the behavior of its inhabitants as completely as possible. The programs, which consist of a great variety of procedures, are named for the tokens mediating between desired behavior and reinforcers. Because there is such a variety of generally applicable token economies (they have been used to treat psychiatric patients, retarded patients, schoolchildren from all age groups, criminals, problem teenagers, drug addicts, and alcoholics), and because they have proved to be so effective, they have greatly contributed to making the vast potential of behavior modification techniques widely known (cf. Kazdin, 1977, 1983).

In Anna, Azrin and Ayllon were given a group of completely hopeless, chronically psychotic patients and a ward specially equipped for research purposes. The object of the program was to keep the patients busy all day with a variety of useful activities. Some of the procedures adopted by the two researchers may seem astonishingly simple and pragmatic. When one looks at them more closely, however, one realizes how much thought and imagination went into working them out. All ward chores, even some of those that had regularly been done by staff, were classified in order of difficulty and then allocated to patients according to their individual abilities. Activities that the patients particularly enjoyed or engaged in often were used as "back-up reinforcers." These activities, ranging from the right to choose one's bedroom to counseling time with ward staff or even watching television for specified periods, could be exchanged for tokens (cf. Azrin & Ayllon, 1968).

As is the case with programmed learning, what is especially attractive about token economies is, presumably, that they can increase the therapeutic effectiveness of institutions with financial and personnel problems without necessitating changes respecting either one. Moreover, token economies are effective even when other therapy programs have failed. They have succeeded not only in penetrating the apathy of many chronically psychotic patients but also in permitting the permanent reintegration of such patients into society.

Programs like those conducted by Azrin and Ayllon in Anna or Atthowe and Krasner in Palo Alto (Atthowe & Krasner, 1968) aimed primarily at implementing and improving new behavior modification techniques for use in clinics. Second-generation token economy programs, on the other hand, like that carried out by Paul and Lentz (1977), were concerned with issues like the comparative effectiveness of treatment and maintaining and transferring gains achieved in therapy to external settings. When, at the end of their six-year project, traditional psychiatric treatment was compared to milieu therapy and social learning therapy—a token economy program that drew on experience gained in Anna and Palo Alto—Paul and Lentz could point to remarkable success with patients who had been classified as highly disturbed: 10.7% of the learning-therapy patients, 7% of the milieu-therapy patients, but none of the traditionally treated patients left the clinic in the course of the program to lead normal lives without requiring rehospitalization. A total of 25% of the learning-therapy patients and 14% of the milieu-therapy patients no longer differed at all from other normal people. The learning-therapy program proved to be noticeably superior to the milieu-therapy program in all tested phases. Nearly all remaining learning-therapy patients could be discharged into residential care units under the

auspices of a regional psychiatric service project. Before treatment, outpatient status would not have been possible (Paul & Lentz, 1977).

Once outside the clinic, behavior modifiers were powerless to act. Orientation and transference procedures had indeed been developed quite early along the lines of social-learning principles. But factors like negative attitudes to the mentally ill, regional infrastructure, and inadequate outpatient care actually determine rehabilitation success. Finding regular volunteers to facilitate the reintegration of discharged patients into community life was often impossible. In addition, patients had generally too little education to get work in regions with relatively high unemployment rates. Outpatient counseling units to which patients were assigned tended to be unreliable and inadequate. Therefore, although 25% of the learning-therapy patients were indistinguishable from other normal people by the end of Paul and Lentz's program, only 10.7% could be fully reintegrated into the community on a permanent basis (see Paul & Lentz, 1977).

Token economies are a subtle imitation of the basic structure of modern achievement-oriented societies (see Kagel & Winkler, 1972). Yet no other course of treatment manipulates patients' lives to such an extent. Even when Ayllon was conducting his first token economy experiments on chronically psychotic patients in Saskatchewan Hospital, Wayburn, Canada, the authorities became so concerned about the ethics of the procedures involved that he was not permitted to carry out his research project beyond the first phase (Ayllon & Haughton, 1962; Ayllon & Michael, 1959; cf. Schorr, 1984a). Although later program procedures have been less rigorous, misgivings about the ethics of token economy programs persist. First, to ensure program effectiveness reinforcement withdrawal must work; for token reinforcement to be effective, patients must have experienced withdrawal. Second, there is always the danger that token reinforcement may increase patients' compliance with institutional and social standards instead of leading to personal development (Ulrich & Mueller, 1977).

Systematic withdrawal of reinforcement is particularly open to criticism when punishment procedures such as time out or response-cost mechanisms are built into a program. It is generally accepted that punishment can have the undesired side effect of renewed aggressive behavior (Bandura, 1965; Hutchinson, 1977; Matson & Kazdin, 1981; Skinner, 1953). Unfortunately, this leads to the dilemma that rejecting punishment often reduces effectiveness in both token economy and operant therapy programs. In Paul and Lentz's project (1977), for instance, the strong urge to conformity which was induced in patients in both milieu and learning-therapy programs led to an increase in assaults and other aggressive behavior, initially such patients' main therapy problem. To reduce undesired behavior levels Paul and Lentz took a 72-hour time out and added a response-cost procedure to the learning-therapy program. When a state directive reduced time-out periods from 72 hours to two hours after the project had been going on for some years, both therapy groups recorded serious disruption that in some cases made a considerable difference in treatment effectiveness. Not until the directive had been rescinded and the former time-out periods had been reinstated could the program regain its earlier effectiveness (Paul & Lentz, 1977).

During the 1970s the use of time-out procedures was sharply limited in the United States by court order. Isolating a patient for treatment or rehabilitation over longer periods is no longer allowed. For short isolation periods the patient must at all times be able to satisfy basic needs (Martin, 1975; Stolz, 1978a). Most ethical and legal problems arose in connection with operant programs treating patients who had been hospitalized involuntarily (for example, prisoners and psychiatric patients). Difficulties in some cases resulted from extremely restrictive and even abusive implementation of programs such as the token economy program at Patton State Hospital (Wexler, 1975) and the START Program at the Medical Center for Federal Prisoners in Springfield (Stolz, 1978a). House review committees and regulations like those developed by various scientific and professional associations (APA, AABT) as a response to what happened there are designed to prevent abuses of behavior therapy procedures. Matson and Kazdin (1981) point out that the use of aversive therapy must be preceded by painstaking cost-benefit analysis if no viable alternatives are available. Treatment goals must be closely scrutinized. This holds not only for aversive elements but also for operant-therapy programs in general. As Kazdin said in 1977 about the objectives of token economy programs: "It is not clear what the realistic and ideal treatment goals are with diverse populations" (1977, p. 280). He went on to add, "Actually the goals need to be individualized across populations and, of course, even within populations."

Hitches in technological planning such as the emergence of "nonresponders" in token economies show the dimensions of the problem. As is still the case, there were patients in early programs who did not respond to treatment (Azrin & Ayllon, 1968: 18%; Atthowe & Krasner, 1968: 10%). For specialists like Kazdin (1983), turning nonresponders into "responders" is primarily a matter of using contingencies in a sophisticated manner. Hersen may well be right in questioning whether a purely technological solution to the problem is adequate. In an article on the problems and limitations of behavior therapy in psychiatry, Hersen (1979) describes the following case:

I was asked to interview and help devise a program for a 65-year-old patient (diagnosis—simple schizophrenia) with an extensive history of brief hospitalizations (each of which was usually preceded by arrest for vagrancy) in many states for over a 40-year period. This patient was residing on a locked ward under token economy lines, but he was described as a nonresponder. While interviewing this patient, it became crystal clear that he had been living the life of a hobo (i.e., riding the freights, working when he required money, etc.), and that he thoroughly enjoyed his life, expressing absolutely no regrets about it. It is little wonder, then, that he was a nonresponder to the token system, particularly as he also was a nonresponder to the large-scale token economy (middle-class America). Indeed, the ultimate "turn on" (reinforcer) for this patient was to roam the country as he had done so for 40 odd years. Given his "lust" for freedom, it obviously was wrong, behaviorally and existentially, to attempt to get him to conform within the confines of the ward token economy. Therefore, I recommended that he be discharged, inasmuch as he certainly was of no danger to others or himself. (pp. 72-73)

Behavior therapists have not yet worked out how to take into consideration the ethical, legal, and social aspects of therapy goals. Finally, the unsystematic and erratic way in which therapy goals are projected as well as the strong technological bias of

program development have promoted uncritical and incorrect implementation of token economy programs. Experience gained in the 1970s has apparently done little toward encouraging efforts to counter this fundamental deficiency. The importance of this question becomes clear when one realizes that developing token economy programs in the operant paradigm was not an isolated achievement but a basic element of a wider program. Token economies were the first significant step toward applying the principles of operant conditioning to social engineering (Ulrich & Mueller, 1977). In articles like "The Design of Cultures" (1961), Skinner makes an earnest plea for the use of his learning theory in social engineering. His utopian novel, *Walden II* (Skinner, 1948), made an early case for its social engineering potential. In *Beyond Freedom and Dignity* (1971), Skinner contends that modern societies cannot survive without scientific planning. A much more broadly based application of operant behavior modification to the daily lives of normal human beings had become the logical consequence of the operant paradigm. Then severe public criticism of behavior modification programs in prisons and psychiatric clinics had the—from the clinical standpoint—unfortunate effect of discouraging applied behavior therapists from continuing with programs or implementing new ones (cf. Schorr, 1984a). At the same time, however, the trend to apply operant technology to behavior modification in the community has grown more pronounced. There has been less public objection to such programs since participation has been voluntary (Kazdin, 1977, 1984).

There is a discrepancy between the lofty claims made for the operant paradigm in regard to planning and coordinating social systems, on the one hand, and the lip service paid to the ethical, legal, and social aspects of social engineering goals, on the other. The uses to which operant learning theory can be put are represented pragmatically as harmless if they are to promote traditional social goals (Kazdin, 1984). Thus, society, and not the applied researcher, is held responsible for what is done. The inability of Skinner's disciples to realize what is at stake is painfully obvious from Wolf's article "Social Validity: The Case of Subjective Measurement, or How Applied Behavior Analysis Is Finding Its Heart" (1978). Wolf reports:

Almost a decade ago, when the field of applied behavior analysis was beginning to expand so rapidly, we were faced with the task of putting together the *Journal of Applied Behavior Analysis*. For a period of several months Garth Hopkins, who was our managing editor, presented us with a series of unexpected dedications to make. Just a couple of days before we were scheduled to go to press with our very first issue, Garth called with one more question "What is the purpose of the *Journal of Applied Behavior Analysis*?" he asked. He said we needed to put a description of the purpose on the inside cover, as one finds in other journals. He needed the answer almost immediately. (p. 203)

Baer wrote the following description in a great hurry:

The *Journal of Applied Behavior Analysis* is primarily for the original publication of reports of experimental research involving applications of the analysis of behavior to problems of social importance. (Cf. *JABA*, front cover)

A later generation of applied behavior analysts was less worried about actually implementing the realistic but vaguely outlined goal of dealing with significant social issues. The basic question was how to demonstrate the social relevance of objectives, procedures, and results in methodologically acceptable form. As Wolf puts it:

If our objective was, as described in *JABA*, to do something of social importance, then we needed to develop better systems and measures for asking society whether we were accomplishing this objective. (1978, p. 207)

Justifying the required use of subjective data collected from interviews, questionnaires, diagnostic profiles, rating scales, and so on is still a highly controversial but crucial issue.

Holland (1978) deplores the reluctance of colleagues to think about the social implications of what they are doing:

Most behavior modification programs merely arrange special contingencies in a special environment to eliminate the "problem" behavior. Even when the problem behavior is as wide-spread as alcoholism and crime, behavior modifiers focus on "fixing" the alcoholic and the criminal, not on changing the societal contingencies that prevail outside the therapeutic environment and continue to produce alcoholics and criminals. (p. 163)

Stolz (1978b) also suggests that applied behavior analysts should be more aware of the ethical implications of therapeutic goals. She gives a detailed analysis of how reinforcement affects therapists themselves:

Persons involved in the decision-making process should attempt to specify what their values are, what reinforcers are controlling their behavior in the short and long run, and should be sensitive to potential conflicts among the value systems of the true client, the professional and the target individual. (p. 699)

In an historical survey Napoli (1981) called the applied psychologists of the 1930s, 1940s, and 1950s "architects of adjustment." Today's operant behavior therapists, however, should not feel obliged to play this role even more efficiently by applying much more impressive technology. The aim to broaden application to social problems in a wider context could actually intensify conflicting attitudes to ethical and social responsibilities: "Applications to alter social problems at the level of the city, community, state and country may raise concerns about attempts to control behavior," as Kazdin puts it at the end of his book on the uses of behavior modification (1984, p. 298). In a society that has gone through the Holocaust, that protects individual rights, and has a free press, operant technologies can be accepted fully only if their advocates can convey their awareness of social responsibilities in a convincing manner.

CONCLUSION

A glance through the *JABA* reveals the enormous variety of uses to which behavior modification technology has been put by operant behavior therapists in the past decade. The methodology used by the authors of the *JABA* especially differs from that presented in other behavior therapy publications, although this cannot really be said of the behavior change techniques they employ. Conceptual eclecticism is what has made the development of practicable technologies possible. However, the evidence of technological maturity is not yet fully conclusive. Behavior therapists of all schools agree that the most glaring technological deficiency of behavior therapy

lies in the failure to maintain therapy gains and transfer them to normal settings (see Agras & Berkowitz, 1980; Baer, 1982; Kazdin, 1984; Stuart, 1982).

Applied behavior analysts and those of other schools of behavior therapy do not merely have in common several positive attitudes such as the rejection of psychodynamic or quasi-disease models of mental disease and a commitment to empirical methodology (Kazdin & Wilson, 1978). On the negative side, they have failed to work on the theoretical foundations of their technologies (cf. Ross, 1985). Operant behavior therapists are well grounded in experimental method. But attempting to meet all clinical needs with method alone can at best be only a short-term perspective for applied research. A long-term goal could be putting the methodological potential of applied behavior analysis to greater use in expanding the conceptual framework of the whole behavior therapy field.

Up to now, a purely utilitarian bias has determined the implementation of operant technologies. Therefore applied behavior analysts regard themselves as "agents of social control" instead of "agents of social change" (Stolz, 1978b). Behaviorist ethics evaluate behavior as positive if it guarantees society's survival (cf. Schorr, 1985). Society alone as the source of all reinforcers determines which behavior must be altered, or, as Krasner and Ullmann put it in 1965, "The ultimate source of values is neither the patient's nor the therapist's wishes, but the requirements of the society in which both live" (p. 363).

This rather naive outlook appears dangerous when one considers how widely operant technologies are being applied to the behavior of normal people. "Behavior modification models the normal societal control process and makes the process explicit and clearer," concludes Holland (1978, p. 171). Although he is aware of the potential of operant technologies in the hands of committed social reformers, one should pause to consider what effect such a one-sided emphasis on social behavior could have on the individual. Restructuring the individual's social environment by means of operant reinforcement makes self-interest the basis of human relationships. A maximum of reinforcement is guaranteed in the hope that the subject may orient himself to the reinforcement available in society and thus be capable of controlling his environment. At this point one begins to wonder whether the awareness gained by the individual of the reinforcement structures his "real" environment contains may not lead to a feeling that conforming to society's values and norms is futile. Social groups develop characteristics by aspiring to more than purely utilitarian goals. Applying operant technologies uncritically to normal settings could well undermine the underlying set of convictions usually known as "basic ethics." Meichenbaum and his colleagues (1968) record a development of this sort. They had given a group of institutionalized young women contingent reinforcement for correct classroom behavior in the afternoon. The measure was immediately effective, but by the following morning classroom behavior had deteriorated, as if the message were "If you don't pay us, we won't shape up" (Meichenbaum, Bowers, & Ross, 1968, p. 349). The long-term effects of manipulating everyday environments by means of behavior modification programs are not yet known. This also holds for the effects of inducing a predominantly utilitarian outlook in groups of subjects controlled by such programs. We do know, however, that none of the behavior problems approached by means of operant

techniques can be considered permanent or even long-term “cures” (Ulrich & Mueller, 1977). This is a valid argument against the uncritical application of operant behavior modification programs on a larger scale, inasmuch as those responsible for such programs are allowed to establish program goals only on the basis of pseudo-democratic voting systems. Operant behavior analysts should submit society’s values and norms to a regular reevaluation. At the same time a rigorous examination of existing social structures—and not just reinforcement contingencies—that ultimately determine the success of behavior therapy programs must be a long-term goal.

REFERENCES

- Agras, W. S., & Berkowitz, R. (1980). Clinical research in behavior therapy: Halfway there? *Behavior Therapy, 2*, 472–487.
- Atthowe, J. M., & Krasner, L. (1968). Preliminary report on the application of contingent reinforcement procedures (token economy) on a “chronic” psychiatric ward. *Journal of Abnormal Psychology, 73*, 37–43.
- Ayllon, T., & Houghton, E. (1962). Control of the behavior of schizophrenic patients by food. *Journal of the Experimental Analysis of Behavior, 5*, 343–352.
- Ayllon, T., & Michael, J. (1959). The psychiatric nurse as a behavioral engineer. *Journal of the Experimental Analysis of Behavior, 2*, 323–334.
- Azrin, N. H. (1977). A strategy for applied research: Learning based but outcome oriented. *American Psychologist, 32*, 140–149.
- Azrin, N. H., & Ayllon, T. (1968). *The token economy. A motivational system for therapy and rehabilitation*. New York: Appleton-Century-Crofts.
- Baer, D. M. (1982). The role of current pragmatics in the future analysis of generalization technology. In R. B. Stuart (Ed.), *Adherence, compliance and generalization in behavioral medicine*. New York: Brunner/Mazel.
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*, 91–97.
- Bandura, A. (1965). Influence of model’s reinforcement contingencies on the acquisition of imitative responses. *Journal of Personality and Social Psychology, 1*, 589–595.
- Biglan, A., & Kass, D. J. (1977). The empirical nature of behavior therapies. *Behaviorism, 5*, 1–15.
- Deitz, S. M. (1978). Current status of applied behavior analysis. Science versus technology. *American Psychologist, 33*, 805–814.
- Ferster, C. B. (1978). Is operant conditioning getting bored with behavior? *Journal of the Experimental Analysis of Behavior, 29*, 347–349.
- Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.
- Foa, E. B., & Emmelkamp, P. M. G. (Eds.). (1983). *Failures in behavior therapy*. New York: Wiley.
- Franzini, L. R., & Tilker, H. A. (1972). On the terminological confusion between behavior therapy and behavior modification. *Behavior Therapy, 3*, 279–282.
- Harzem, P. (1984). Experimental analysis of individual differences and personality. *Journal of the Experimental Analysis of Behavior, 42*, 385–395.
- Hayes, S. C., Rincover, A., & Solnick, J. V. (1980). The technical drift of applied behavior analysis. *Journal of Applied Behavior Analysis, 13*, 275–285.
- Hersen, M. (1979). Limitations and problems in the clinical application of behavioral techniques in psychiatric settings. *Behavior Therapy, 10*, 65–80.
- Holland, J. G. (1978). Behaviorism: Part of the problem or part of the solution? *Journal of Applied Behavior Analysis, 11*, 163–174.
- Honig, W. K., & Staddon, J. E. R. (Eds.). (1977). *Handbook of operant behavior*. Englewood Cliffs, NJ: Prentice-Hall.

- Hutchinson, R. R. (1977). By-products of aversive control. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Kagel, J. H., & Winkler, R. C. (1972). Behavioral economics: Areas of cooperative research between economics and applied behavior analysis. *Journal of Applied Behavior Analysis*, 5, 335-342.
- Kazdin, A. E. (1977). *The token economy. A review and evaluation*. New York: Plenum Press.
- Kazdin, A. E. (1978). *History of behavior modification. Experimental foundations of contemporary research*. Baltimore: University Park Press.
- Kazdin, A. E. (1983). Failure of persons to respond to the token economy. In E. B. Foa & P. M. G. Emmelkamp (Eds.), *Failures in behavior therapy*. New York: Wiley.
- Kazdin, A. E. (1984). *Behavior modification in applied settings*. (3rd ed.). Homewood, IL: Dorsey Press.
- Kazdin, A. E., & Wilson, G. T. (1978). *Evaluation of behavior therapy. Issues, evidence and research strategies*. Cambridge, MA: Ballinger.
- Keehn, J. D., & Webster, C. D. (1969). Behavior therapy and behavior modification. *Canadian Psychologist*, 10, 68-73.
- Krantz, D. L. (1971). The separate worlds of operant and non-operant psychology. *Journal of Applied Behavior Analysis*, 4, 61-70.
- Krasner, L., & Ullmann, L. P. (Eds.). (1965). *Research in behavior modification. New developments and implications*. New York: Holt, Rhinehart, & Winston.
- Krasner, L. (1971). Behavior therapy. *Annual Review of Psychology*, 22, 483-532.
- Krasner, L. (1980). *Environmental design and human behavior. A psychology of the individual in society*. New York: Pergamon Press.
- Kuhn, T. S. (1976). *Die Struktur wissenschaftlicher Revolutionen*. Frankfurt: Suhrkamp Verlag.
- Kuhn, T. S. (1977). *The essential tension: Selected studies in scientific tradition*. Chicago: University of Chicago Press.
- Lazarus, A. (1971). *Behavior therapy and beyond*. New York: McGraw-Hill.
- Ledwidge, B. (1978). Cognitive behavior modification: A step in the wrong direction? *Psychological Bulletin*, 85, 353-375.
- Ledwidge, B. (1979). Cognitive behavior modification or new ways to change minds: Reply to Mahoney and Kazdin. *Psychological Bulletin*, 86, 1050-1053.
- Mahoney, M. J., Kazdin, A. E., & Lesswing, N. J. (1974). Behavior modification: Delusion or deliverance? *Annual Review of Behavior Therapy*, 2, 11-40.
- Mahoney, M. J., & Arnkoff, D. B. (1978). Cognitive and self-control therapies. In S. Garfield & A. Bergin (Eds.), *Handbook of psychotherapy and behavior change*. New York: Wiley.
- Mahoney, M. J., & Kazdin, A. E. (1979). Cognitive behavior modification: Misconceptions and premature evaluation. *Psychological Bulletin*, 86, 1044-1049.
- Marr, M. J. (1984). Conceptual approaches and issues. *Journal of the Experimental Analysis of Behavior*, 42, 353-362.
- Martin, R. (1975). Legal challenges to behavior modification: Trends in schools, corrections and mental health. Champaign, IL: Research Press.
- Matson, J. L., & Kazdin, A. E. (1981). Punishment in behavior modification: Pragmatic, ethical, and legal issues. *Clinical Psychology Review*, 1, 197-210.
- Meichenbaum, D. H., Bowers, K., & Ross, R. R. (1968). Modification of classroom behavior of institutionalized female adolescent offenders. *Behavior Research and Therapy*, 6, 343-353.
- Napoli, D. S. (1981). *The architects of adjustment. The history of the psychological profession in the United States*. Port Washington, NY: Kennikat Press.
- Neuringer, A. (1984). Melioration and self-experimentation. *Journal of the Experimental Analysis of Behavior*, 42, 397-406.
- Nevin, J. A. (1980). Editorial. *Journal of the Experimental Analysis of Behavior*, 33.
- Nietzel, M. T., Winett, R. A., MacDonald, M., & Davidson, W. S. (1977). *Behavioral approaches to community psychology*. Elmsford, NY: Pergamon Press.
- Paul, G. L., & Lentz, R. J. (1977). Psychosocial treatment of chronic mental patients. Milieu versus social-learning programs. Cambridge, MA: Harvard University Press.
- Pomerleau, O. F. (1979). Behavioral medicine. The contribution of the experimental analysis of behavior to medical care. *American Psychologist*, 34, 654-663.

- Reynolds, G. S. (1975). *The primer of operant conditioning*. Glenview, IL: Scott Foresman.
- Risley, T. R. (1970). Behavior modification. An experimental-therapeutic endeavour. In L. A. Hamerlynck, P. O. Davidson, & L. E. Acker (Eds.), *Behavior modification and ideal mental health services*. Calgary, Alberta: University of Calgary Press.
- Ross, A. O. (1985). To form a more perfect union: It is time to stop standing still! *Behavior Therapy*, *16*, 195–204.
- Schorr, A. (1984a). *Die Verhaltenstherapie. Ihre Geschichte von den Anfängen bis zur Gegenwart*. Weinheim: Beltz-Verlag.
- Schorr, A. (1984b). Clark Hull und die Psychoanalyse. In S. Bem, H. Rappard, & W. van Hoorn (Eds.), *Studies in the History of Psychology and the Social Sciences* (vol. 2). Leiden: Psychologisch Instituut van de Rijksuniversiteit.
- Schorr, A. (1985). Behaviorismus, Ethik und die utopische Gesellschaft. In S. Bem, H. Rappard, & W. van Hoorn (Eds.), *Studies in the History of Psychology and the Social Sciences* (Vol. 3). Leiden: Psychologisch Instituut van de Rijksuniversiteit.
- Shimp, C. P. (1984). Cognition, behavior, and the experimental analysis of behavior. *Journal of the Experimental Analysis of Behavior*, *42*, 407–420.
- Skinner, B. F. (1935). Two types of conditioned reflex and a pseudo-type. *Journal of General Psychology*, *12*, 66–77.
- Skinner, B. F. (1937). Two types of conditioned reflex: A reply to Konorski and Miller. *Journal of General Psychology*, *16*, 272–279.
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century.
- Skinner, B. F. (1945). The operational analysis of psychological types. *Psychological Review*, *52*, 270–277.
- Skinner, B. F. (1948). *Walden Two*. New York: Macmillan.
- Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review*, *57*, 193–216.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
- Skinner, B. F. (1961). The design of cultures. *Daedalus*, *90*, 534–546.
- Skinner, B. F. (1970). What is the experimental analysis of human behavior? In R. Ulrich, T. Stachnik, & J. Mabry (Eds.), *Control of human behavior* (Vol. 2). Glenview, IL: Scott, Foresman.
- Skinner, B. F. (1971). *Beyond freedom and dignity*. New York: Knopf.
- Skinner, B. F. (1974). *About behaviorism*. New York: Knopf.
- Skinner, B. F. (1977). Why I am not a cognitive psychologist. *Behaviorism*, *5*, 1–10.
- Skinner, B. F. (1986). The evolution of verbal behavior. *Journal of the Experimental Analysis of Behavior*, *45*, 115–122.
- Stolz, S. B. (Ed.). (1978a). *Ethical issues in behavior modification*. Report to the American Psychological Association Commission. San Francisco: Jossey-Bass.
- Stolz, S. B. (1978b). Ethics of social and educational interventions: Historical context and a behavioral analysis. In A. C. Catania & T. A. Brigham (Eds.), *Handbook of applied behavior analysis: Social and instructional processes*. New York: Irvington.
- Stuart, R. B. (Ed.). (1982). *Adherence, compliance and generalization in behavioral medicine*. New York: Brunner/Mazel.
- Sundel, M., & Sundel, S. S. (1982). *Behavioral medicine in the human services: A systematic introduction to concepts and applications*. Englewood Cliffs, NJ: Prentice-Hall.
- Ulrich, R. E., & Mueller, K. (1977). Operante Kontrolle von Sozialverhalten und Gesellschaft. In H. Zeier (Hg.), *Psychologie des zwanzigsten Jahrhunderts. Band VI: Pawlow und die Folgen*. München: Kindler Verlag.
- Wexler, T. (1975). Token and taboo: Behavior modification, token economy, and the law. In C. M. Franks & G. T. Wilson (Eds.), *Annual Review of Behavior Therapy* (Vol. 2). New York: Brunner/Mazel.
- Wolf, M. M. (1978). Social validity: The case for subjective measurement or, How applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, *11*, 203–214.
- Zuriff, G. E. (1985). *Behaviorism: A conceptual reconstruction*. New York: Columbia University Press.

PART II

CONDITIONING THEORY

CHAPTER 3

Animal Conditioning and Learning Theory

Anthony Dickinson

The original premise of behavior therapy was that certain pathological behavior patterns are acquired through conditioning and therefore treatable by controlled and appropriate manipulation of the processes underlying this form of learning. This assumption places conditioning at the theoretical focus of any discussion of abnormal behavior, not only for those who endorse the premise but also for those who wish to challenge it. And for both parties the analysis and treatment of such disorders must be measured against contemporary views of conditioning rather than those current at the genesis of behavior therapy a generation or so ago. This would be of little import if our view of conditioning had remained relatively static; the fact, however, is that conditioning theory has undergone a major revision during the intervening years.

In the 1940s and 1950s the study of conditioning was dominated by neobehaviorism. The job of psychology in general was seen to be that of understanding and predicting behavior, with learning theory in particular being assigned the problem of acquired behavior. This perspective elevated conditioning to the central paradigm for studying learning, for in conditioning we observe directly the acquisition and strengthening of a new behavior. Moreover, learning appeared to be amenable to an analysis in simple behavioral terms without appealing to any mental entities or processes. Conditioning just involved the strengthening or reinforcement of a response brought about by the presentation of a reinforcing agent immediately after the response or its eliciting stimulus. For the behaviorist learning and reinforcement were one and the same thing.

Over the last 20 years or so, almost every aspect of this view of conditioning has been challenged. We have found that even the simplest forms of conditioning appear to involve cognitive processes, that learning can occur without reinforcement within the conditioning paradigm, and that even where reinforcement operates, strict

contiguity between the response or stimulus and the reinforcer is neither necessary nor sufficient for conditioning. In this chapter I shall attempt to document the empirical evidence underlying these claims and, in doing so, I hope to characterize the contemporary view of conditioning, at least as seen by a learning theorist.

I shall start out by describing a couple of phenomena that, I think, require us to take a cognitive view of conditioning before discussing the major factors affecting the acquisition of conditioning that have been discovered during the last two decades or so. These discoveries have spawned a number of contemporary theories of conditioning that will be discussed before I turn finally to a couple of special topics: inhibitory conditioning and the recent discovery of the role of conditional associative relationships that appears to transcend simple conditioning. My discussion will be restricted to animal conditioning, for all the major theoretical revisions have arisen from research in this area.

COGNITION AND CONDITIONING

The cognitive view of conditioning claims that the development of the conditional response reflects the acquisition of knowledge about the relationship between the events in a conditioning experience rather than the direct strengthening of a response or behavioral disposition posited by reinforcement theory. To illustrate the motivation for this claim, I shall start out by considering a phenomenon, reinforcer reevaluation, that is not readily amenable to a reinforcement analysis, but makes direct contact with a knowledge-based account. I shall then discuss a study that demonstrates that we can also analyze the form of the knowledge underlying conditioning. Knowledge about event or stimulus relationships must involve some form of internal or mental representation of these stimuli. The second phenomenon, mediated conditioning, provides a fairly direct demonstration of the operation of such representations in conditioning.

REINFORCER REEVALUATION

The contrast between reinforcement and cognitive approaches can be illustrated by considering the case of instrumental conditioning. In a simple instrumental procedure the performance of an action, such as pressing a lever in an operant chamber, causes the delivery of a positive reinforcer or reward, such as a food pellet. As a result of arranging this contingency, not surprisingly, the animal comes to perform the instrumental action more frequently, and this form of conditioning is often taken as a paradigmatic example of how we control our environment to achieve desirable goals.

Despite the apparent simplicity of the instrumental procedure, classic neobehaviorists, such as Hull and Guthrie, and cognitive psychologists offer contrasting accounts of the learning that underlies this form of conditioning. For the behaviorist, presenting the reinforcer immediately after a lever press in the presence of some stimulus, such as the sight of the lever, strengthens the capacity of that stimulus to elicit the response on subsequent occasions. Such strengthening is, of course, the process of reinforcement. It is important to note that this account makes no appeal

to any cognitive or mental entities and processes; the animal does not acquire any knowledge about the relationship between pressing the lever and the delivery of food, but rather the role of the reinforcer is just to increase the tendency to make the response to the appropriate stimulus. This view is in marked contrast to that of the cognitive and lay psychologist, who would both argue that the hungry animal lever presses because it knows that this action causes the delivery of food.

The distinction between the behavioral and knowledge-based interpretations of conditioning has long been recognized, as has the nature of the empirical test to distinguish between them (e.g., Tolman, 1933), the reinforcer revaluation procedure. Suppose that having trained the animal to press the lever for a particular type of food, we independently devalue this food so that the animal will no longer eat it even when hungry. We can now ask what the animal will do when it is once again given the opportunity to lever press. If the animal knows that lever pressing causes the delivery of this particular food, it should be reluctant to press following reinforcer devaluation relative to another animal that has not developed an aversion to the food. This should be true even though lever pressing is tested in extinction in absence of any presentations of the food. Testing in extinction is important because it ensures that any differences in performance must be mediated by the stored knowledge of the instrumental contingency rather than the direct effect of experiencing the devalued food on performance. The problem with this test is that until the relatively recent development of food-aversions procedures, there was no effective method of devaluing a positive reinforcer without changing the animal's motivational state.

Adams and I (Adams & Dickinson, 1981) have studied the effect of reinforcer revaluation by food-aversion training on instrumental performance. The basic idea was to train the animals with two potential reinforcers, only one of which was contingent upon their action, lever pressing. We could then compare the effect of devaluing the contingent and noncontingent reinforcers. The design of the study is illustrated in Table 1. On alternate days (days n) we trained rats to lever press for one type of food pellet, the reinforcer, whereas on the other days (days $n + 1$) we simply placed the animals in the operant chamber without the lever being present and gave them free access to the same number of another type of pellet, the noncontingent food. For half the animals the reinforcer was sucrose pellets (Suc) and the noncontingent food mixed-diet pellets (Pel) with the remaining rats receiving the opposite assignment. After performance had been established, the animals continued to receive the two food types on alternate days but now both of them were presented

TABLE 1. Design of a Reinforcement Devaluation Experiment^a

Group	Training		Devaluation	
	Day n	Day $n + 1$	Day n	Day $n + 1$
Paired	Lp - Suc	Pel	Suc - Li	Pel
	or Lp - Pel	Suc	Pel - Li	Suc
Unpaired	Lp - Suc	Pel	Pel - Li	Suc
	or Lp - Pel	Suc	Suc - Li	Pel

^aFrom Adams and Dickinson (1981). Lp = lever press; Suc = sucrose reinforcer; Pel = pellet reinforcer; Li = lithium chloride.

freely in the absence of the lever. In addition, the animals were injected with an emetic, lithium chloride (Li), immediately after they consumed one type of food but not the other. The injection followed consumption of the reinforcer for the paired group and consumption of the noncontingent food for the unpaired group. As rats rapidly learn to stop eating a food followed by sickness, this design ensured the devaluation of the reinforcer for paired animals but not for the unpaired subjects, while exposing both groups to the same aversion procedure. When the animals once again had the opportunity to press the lever in the extinction test, we observed a clear devaluation effect; the paired group, which had been averted to the reinforcer, pressed significantly less than the animals in the unpaired or control group.

It is not clear how one could give an account of this finding in terms of simple stimulus-response/reinforcement theory. If all the reinforcer does is to strengthen a stimulus-response association, in this case between the lever and pressing it, changing the value of the reinforcer once that association is formed should have no effect on performance. It is true that more sophisticated, two-factor versions of stimulus-response theory have been developed (e.g., Rescorla & Solomon, 1967; Trapold & Overmier, 1972), but they still have difficulty in accounting for the devaluation effect (see Adams & Dickinson, 1981; Colwill & Rescorla, 1985). There is little reason to resist the obvious conclusion that during instrumental conditioning animals learn about the relationship between their actions and the associated consequences. This knowledge can then be integrated with information about the value of the reinforcer to control performance.

MEDIATED CONDITIONING

To say that an animal knows about the relationship between an action and its reinforcer is to claim that it has some internal representation of the reinforcer that can control behavior in the absence of this stimulus. In giving a cognitive interpretation of the Adams and Dickinson study, we inferred the role for such a representation from the fact that if we changed the value of the actual reinforcer, we observed a change in behavior that was difficult to understand unless we assumed that there was a corresponding alteration in the value associated with some internal representation of this stimulus. The reality of this representation and of the commerce between it and the actual stimulus it represents would be greatly enhanced if we could also demonstrate the opposite transaction, namely a transfer of value from the representation to the actual stimulus. This Holland (1981) did in a recent study of mediated conditioning.

Holland employed a classical or Pavlovian procedure in contrast to the instrumental conditioning studied by Adams and Dickinson. Rather than arranging an instrumental relationship between an action and a reinforcer, the Pavlovian procedure exposes the subject to an association between a stimulus and the reinforcer. So, during the training stage of his experiment, Holland signaled the delivery of food pellets (Pel) by a tone and the presentation of a sucrose solution (Suc) by a light on a number of occasions for two groups of rats (see Table 2). The occurrence of Pavlovian conditioning is shown by the fact that the signaling stimulus, the *conditional stimulus*, comes to elicit a new response, the *conditional response*.

TABLE 2. Design of a Mediated Conditioning Experiment.^a

Group	Training	Revaluation	
		Day <i>n</i>	Day <i>n</i> + 1
Pel	Tone - Pel, Light - Suc	Tone - Li	Light
Suc	Tone - Pel, Light - Suc	Light - Li	Tone

^aFrom Holland (1981). Suc = sucrose reinforcer; Pel = pellet reinforcer; Li = lithium chloride.

The cognitive interpretation of this type of conditioning is that the animals learn about the relationship between the stimulus and its associated reinforcer so that the presentation of the signal now activates or retrieves an internal representation of the reinforcer. Thus, for Holland's rats the tone should have activated a representation of the pellets and the light a representation of the sucrose. Holland attempted to assess this idea in the second stage by pairing a presentation of either the tone or the light with a lithium chloride injection on one set on days (day *n*) and the other stimulus with nothing on the alternate days (day *n* + 1). This means that Group Pel, which received the tone in association with the lithium injection, should also have experienced a pairing of the representation of the pellet with the emetic. Presenting the tone should have activated a representation of the pellet, which was then in a position to be associated with the malaise induced by the lithium. Correspondingly, the pairing of the light with the lithium injection in Group Suc should have produced an aversion to the sucrose solution. To assess whether this was so, Holland tested whether his rats would consume the pellets and sucrose after this conditioning procedure. In accord with cognitive theory, Group Pel ate less pellets than Group Suc, whereas the reverse was true for sucrose consumption. An aversion appeared to have been established to these foods without the animals ever having experienced them in direct association with illness. Thus, it would appear that activating an internal representation of a stimulus can substitute for the actual presentation of that stimulus in simple conditioning.

I have described the reinforcer revaluation and mediated conditioning effects in some detail because it is important to realize that the development of a cognitive perspective is not just a matter of intellectual fashion but has been driven by empirical evidence, of which the two studies I have considered are but examples (see Dickinson, 1980, and Mackintosh, 1983, for a more extensive discussion). Given such effects, it is difficult to escape the conclusion that at least some forms of conditioning reflect the acquisition of knowledge about the relationship or association between the events involved in the conditioning experience rather than the simple strengthening of a stimulus-response link through a reinforcement process.

BEHAVIORAL AUTONOMY

It is rare that a psychological theory is completely wrong; more typically, the problem is that its scope has been overextended. So it is with stimulus-response theory, for there is clearly some truth in it. The fact that we often find ourselves

persisting in instrumental behavior that is no longer appropriate to our current goals suggests the operation of a stimulus-response mechanism. We usually refer to such responses as habits. It is as though the control of the action has become independent of our knowledge about its consequences and, as a result, autonomous of the current value of the goal or reinforcer. This form of behavioral autonomy can be demonstrated using the reinforcement revaluation procedure. For instance, I and my colleagues (Dickinson, 1985) have found that if the instrumental response is trained on certain schedules of reinforcement or the training is extended, devaluing the reinforcer by a food-aversion technique, similar to that used by Adams and Dickinson (1981), has no effect on subsequent responding in an extinction test. It is as though with overtraining there is a shift in control from a cognitive to a stimulus-response mode.

A form of behavioral autonomy can also be demonstrated in Pavlovian conditioning. Although the conditional response is usually reduced in an extinction test by devaluation of the reinforcer, second-order responses often appear to be autonomous. In contrast to first-order conditioning in which the signal is paired directly with the reinforcer, second-order conditioning is brought about by the subsequent arrangement of a predictive relationship between a second signal and the first. Conditioning to the second signal is then often found to be autonomous of the status of the training reinforcers in the sense that it survives both the devaluation of the primary reinforcer and the extinction of conditioning to the first-order signal (Rescorla, 1980). Thus, neither stimulus-response nor knowledge-based theories exert a total hegemony over conditioning, and both have their place in the modern analysis of this form of learning.

CONDITIONS OF ACQUISITION

Just as our view of the associative knowledge set up during conditioning has become more sophisticated in the last 20 years, so has our understanding of the factors affecting the acquisition of conditioning. Kimble's (1961) revision of Hilgard and Marquis' classic text on conditioning and learning identified the amount of training, the size of the reinforcer, the schedule of reinforcement, and the temporal relationship in pairings of the stimulus or action with the reinforcer as the major factors in determining the strength of conditioning. In addition, we should now recognize the temporal correlation between events and the relationship between the qualitative properties of the events as being important. These two new factors are more than just simple additions to the list, for their discovery has radically altered our view of the mechanisms underlying conditioning.

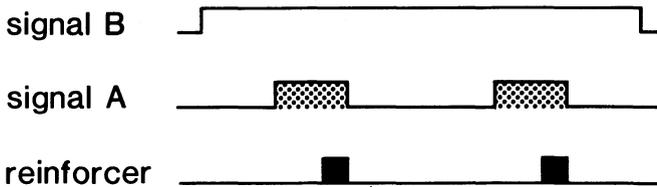
EVENT CORRELATION

Traditional reinforcement theory held a very strong position on the role of the temporal relationship between the signal or action and the reinforcer; contiguous pairings of these events are both necessary and sufficient for conditioning. From a functional point of view, this is in many ways a surprising claim. If instrumental

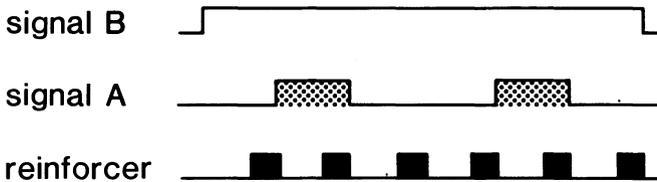
conditioning enables an organism to gain control over its environment, we should expect the learning mechanism underlying conditioning to distinguish real causal relationships between actions and their consequences from fortuitous or accidental associations. And, yet, if this mechanism is sensitive only to temporal pairings, the animal will be prone to numerous superstitious and erroneous beliefs about the causal structure of its environment. The world is full of chance and, often, frequent conjunctions of events that belong to independent causal chains.

This point is illustrated in Figure 1, which represents the temporal sequence of two events, signal A and a reinforcer. In the top sequence signal A is a good candidate as a cause of the reinforcer; the reinforcer consistently follows A and never occurs without it, and this relationship would be detected by a contiguity mechanism simply on the basis of the pairings of the two events. By contrast, the contiguity

Positive Correlation



Zero Correlation



Negative Correlation

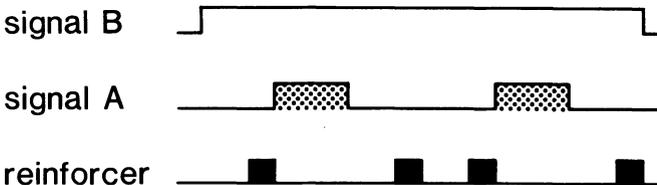


FIGURE 1. A schematic representation of three relationships between signal A and the reinforcer. In the top relationship signal A and the reinforcer are positively correlated, in the middle relationship they are uncorrelated, and in the bottom relationship there is a negative correlation between signal A and the reinforcer. Signal B represents the background or contextual cues.

mechanism would fail as a causality detector in the case of the relationship portrayed in the middle sequence. Here there are just as many pairings as in the first sequence so that the learning mechanism should yield the same relationship in the two cases. And yet we should not attribute a causal status to signal A in the middle sequence; the effect or reinforcer is just as likely to occur in absence of signal A and as in its presence.

What distinguishes the sequences is the correlation or contingency between the two events. In the top sequence there is a positive correlation or contingency between the events, whereas in the second the events are uncorrelated or noncontingent. Rescorla (1968) was the first to investigate systematically whether event correlations affect conditioning, in this case between a signal and a reinforcer in a Pavlovian procedure. He exposed one group of animals to a sequence corresponding to the positive correlation in Figure 1 and another group to the uncorrelated sequence. Although the correlated group showed good conditioning, the signal was totally ineffective following training with a zero correlation in spite of the numerous pairings of the signal and reinforcer experienced by these animals. These two conditions are but a subset of those over which Rescorla demonstrated that the strength of conditioning is systematically related to event contingencies.

Thus, it appears that temporal contiguity is not sufficient for conditioning, but only operates to produce sustained conditioning when embedded within a positive correlation between the signal and reinforcer. How is it, though, that conditioning can track the correlation between events?

SELECTIVE CONDITIONING AND BLOCKING

The initial reaction to Rescorla's (1968) findings was to suppose that animals acquired direct knowledge of the contingency or correlation (see Hammond & Paynter, 1983, for a contemporary discussion of this idea). Rescorla and Wagner (1972), however, have offered an explanation that does not depend on such knowledge, but rather appeals to the psychologically simpler idea of selective conditioning. They pointed out that we never experience a relationship between two events in a vacuum; there are always numerous other potential signals and causes of the reinforcer present in the environment, such as the background cues provided by the context in which conditioning takes place. These background cues are represented by signal B in Figure 1. Seen within this framework, the problem facing the animal is to decide whether the occurrence of the reinforcer is related to presentation of the signal A rather than simply being in the conditioning context. Perhaps in the uncorrelated condition the presentation of the reinforcer in the absence of signal A leads the animal to attribute all occurrences of the reinforcer simply to being in this particular context. This attribution may then, in turn, prevent the animal learning about the relationship between signal A and the reinforcer when the two events are paired, so that little or no conditioning develops to this signal.

There is in fact good evidence that an animal shows little conditioning from pairings of a potential signal and a reinforcer if such pairings occur in the presence of another, well-established signal. This is Kamin's (1969) classic "blocking" effect. The basic design of an experiment revealing the blocking effect is shown in Table 3.

TABLE 3. Design of a Blocking Experiment

Group	Stage 1	Stage 2
Control		AB + ^a
Blocking	B +	AB +
Surprise	B -	AB +

^aA = added signal; B = pretrained signal; + = reinforcement; - = nonreinforcement.

In the first stage conditioning is established to signal B by pairing it with the reinforcer in the blocking group but not in the control group. Then, in a second stage a novel stimulus, signal A, is added and the compound of the pretrained signal B and this novel signal A is paired with the reinforcer a number of times. If animals learn less about the added signal when it is paired with the reinforcer in the presence of a well-established predictor of the reinforcer, we should expect to observe less conditioning to signal A in the blocking group than in the control group, which simply receives the compound training. This is just what Kamin (1969) observed when he subsequently tested conditioning to signal A by presenting it alone; the pretrained stimulus blocked the conditioning that would otherwise have accrued to the added stimulus.

Kamin suggested that blocking occurs because only surprising and unexpected reinforcers are capable of supporting sustained conditioning. The animals in the blocking group learned to expect the reinforcer following the pretrained stimulus B in the first stage so that the presentation of the reinforcer following the AB compound in the second stage was fully expected and, hence, would not support conditioning to signal A. Kamin's surprise theory is backed up by a variety of evidence, although perhaps the most compelling comes from a subsequent study by Rescorla (1971). He added a third group to the basic blocking design, the surprise condition illustrated in Table 3. Rescorla argued that if the strength of conditioning depends on how surprising the reinforcer is, we should be able to get superconditioning rather than blocking if we could make the occurrence of the reinforcer on AB compound trials supersurprising during Stage 2. To do this, he presented the animals in a third, surprise group with a negative correlation between the pretrained stimulus B and the reinforcer during Stage 1. These animals received just as many presentations of the reinforcer and stimulus B during Stage 1 as the blocking group but, instead of being paired with this signal, the reinforcers occurred at random points during Stage 1 training except during and immediately following presentations of the stimulus B. Thus, animals in the surprise group may well have learned that signal B predicted the absence of the reinforcer during Stage 1 (see following), so that its occurrence following the AB compound in Stage 2 should have been very surprising. In accord with this analysis, Rescorla found more conditioning to the added stimulus A in the surprise group than in the controls.

In summary, we have seen that the problem of why conditioning is sensitive to the overall correlation between events can be reduced to that of why blocking occurs. In turn, this question can be answered by explaining why only surprising reinforcers produce sustained conditioning. As we shall see, this is the central problem addressed by contemporary theories of conditioning.

LEARNED IRRELEVANCE

Rescorla's (1968) classic contingency study demonstrated that little or no sustained conditioning occurs from experiencing a random or uncorrelated schedule of reinforcement however often the signal and reinforcer are paired. But this should not be taken as evidence that animals fail to learn in the absence of a contingency, and a commonsense view would lead us to expect that if they learn anything, it should be that the two events are unrelated. The problem is that such learning may well be behaviorally silent for there is no obvious reason why it should be manifest in conditioning. Knowing that a stimulus and reinforcer are unrelated has no implications for the appropriate response to that stimulus.

Mackintosh (1973) argued, however, that such knowledge might be revealed by a transfer test. If an animal learns from an uncorrelated schedule that a stimulus and reinforcer are unrelated, it should take longer to learn subsequently that the stimulus is a signal for the reinforcer when the two are presented in a predictive relationship. In accord with this argument, Mackintosh (1973) found that the development of conditioning to a stimulus paired with a reinforcer was retarded if the animals had been preexposed to random presentations of these two events. This retardation was seen in comparison with control subjects that received prior exposure to either the stimulus alone or the reinforcer alone. It was as though the animals learned on the uncorrelated schedule that the two events were irrelevant to each other.

Mackintosh (1973) also observed another effect, although in this case it was one that was already well established, the so-called latent inhibition effect (Lubow, 1973). The group preexposed to the signal alone also showed slower subsequent conditioning when this stimulus was paired with the reinforcer. Animals appear to be capable of a form of learning that effectively encodes the fact that a stimulus is irrelevant, not just with respect to a particular reinforcer, but also as a predictor of anything of significance, although such learning does not retard conditioning as much as exposure to a random relationship between the stimulus and the particular reinforcer used during conditioning.

CAUSAL RELEVANCE

So far we have seen that conditioning is sensitive to the correlation between events. And so it should be; given the correlation is based on sufficient observations, it is a reliable indicator of the presence of a real relationship between the events. But what should an animal make of only one, or even just a few pairings of a signal and reinforcer? If this stimulus and reinforcer have not occurred at other times in the animal's experience, the correlation will be perfect and yet, being based on so few observations, still be an unreliable indicator of the presence of a real relationship between these two events. From the individual animal's point of view, such pairings may well have arisen by chance rather than through the operation of a causal process. Under these circumstances, a conditioning system evolved to detect the causal structure of the environment should have developed a sensitivity to other indicators of the presence of a real relationship.

An obvious indicator is what is often referred to as the qualitative or intrinsic properties of the signal and reinforcer. The importance of such properties can be illustrated by considering an experiment of Domjan and Wilson (1972) that employed a food aversion procedure similar to that used in the reinforcer devaluation study that we considered. They exposed thirsty rats to one of two signals, either a saccharin flavored solution or a salient exteroceptive stimulus, a noise, before administering an emetic to induce nausea as the reinforcer. After three pairings, they measured the extent to which the animals avoided consuming either the flavored solution or water accompanied by the noise. Although both these stimuli were equally correlated with illness, they found a greater reluctance to drink the saccharin than the “noisy water.” This finding can be understood if it is assumed that stronger conditioning accrues from pairings of the signal and reinforcer when the signal is of a type that is likely to be a cause of that reinforcer. A flavored solution is much more likely to cause an illness or at least to be associated with the cause of an illness than is an exteroceptive stimulus, such as the noise.

This account predicts that if Domjan and Wilson had changed the reinforcer to a type that was causally relevant to the noise rather than the flavored solution, they should have reversed the pattern of avoidance. This they did by using a shock rather than the toxin as the reinforcer in a second group of animals. On test these animals avoided the “noisy water” more than they did the saccharin. This finding also makes sense in terms of causal relevance, for an exteroceptive reinforcer, such as peripheral pain, is much more likely to have been caused by an agent, such as a predator, whose presence is indicated by another external signal, a noise in this case, rather than by an agent related to food ingestion.

Although this discussion has concentrated on causal relevance, it is but one of a number of qualitative factors that determine the strength of conditioning; for instance, both signal-reinforcer similarity and spatial contiguity have been shown to exert an effect (e.g., Rescorla, 1980; Testa, 1974).

LONG-DELAY CONDITIONING

Kamin’s blocking effect revealed the inadequacy of a central tenet of traditional reinforcement theory, namely that temporal contiguity between a signal and an effective reinforcer is sufficient for conditioning. We are also in a position to challenge the necessity of temporal contiguity. It has been known for a number of years (Revusky & Garcia, 1970) that food aversions can be conditioned even if a matter of minutes or even hours elapses between the ingestion of the food and the induction of the illness, an interval very much longer than that over which conventional forms of conditioning can be established.

The initial reaction to this finding was that food aversion conditioning depends on some special learning mechanism. But, as Revusky (1971) pointed out, this is unlikely given the numerous similarities between food aversion learning and standard forms of conditioning. Rather, he argued that long-delay learning could be understood in terms of the type of selective conditioning seen in the blocking effect when coupled with the principle of causal relevance. The blocking effect demonstrates that stimuli in some sense compete to be established as signals for the reinforcer. Given this fact,

Revusky noted that no causally relevant competing stimulus (i.e., another flavor) intervenes between the target signal and the illness reinforcer in a food aversion procedure, whereas in the standard exteroceptive conditioning the contextual cues typically provide numerous interpolated, causally relevant, and thus potentially competing stimuli. It is these stimuli, Revusky argued, that limit the temporal interval over which the signal and reinforcer can be associated. In line with this argument, Revusky found that if he gave his rats a second, causally relevant fluid to drink during the interval between consuming the target fluid and the induction of illness, the aversion to the target was markedly reduced.

Thus, long delay learning accords perfectly with a functional perspective that views conditioning as a process that allows animals, and presumably ourselves, to determine the causal structure of the environment. Conditioning appears to be finely tuned to permit us to learn which events are causally related and which are unrelated, thus endowing us with the ability to both predict and control important events. We shall now consider the various theoretical accounts of the processes underlying this ability.

THEORIES OF CONDITIONING

Three major theories of conditioning have been advanced over the last decade and a half: the Rescorla-Wagner theory (1972) and its subsequent development by Wagner (1978, 1981), the Pearce-Hall (1980) model, and Mackintosh's (1975) theory. As none of them have, as yet, achieved clear ascendancy over the others, I shall briefly consider each of these models. Comparison is eased by the fact that they all focus on Pavlovian conditioning, which they view from a common theoretical perspective, according to which learning is an incremental (and possibly, decremental) process affecting the strength of the associative knowledge about the relationship between the signal and reinforcer. This associative knowledge is assumed to map monotonically into conditioning; the higher the associative strength of a stimulus, the stronger will be the conditioned response elicited by that signal. Moreover, all of the theories assume that these increments occur only in learning episodes involving the actual presentation of the stimulus. Given this perspective, the theorists see their job as that of specifying the factors that determine the size of the increments in associative strength accruing from each conditioning episode. The original presentations of all of the theories did this in the form of linear equations, but I shall not attempt to describe this formalism here. Rather, I shall concentrate on describing what I see as the central psychological ideas captured by these equations, albeit in a very sketchy form.

THE WAGNER THEORY

In the discussion of the conditions of acquisition, I attributed the ability to track event correlations to the fact, revealed by the blocking effect, that only surprising or unexpected reinforcers sustain conditioning. This observation lies at the heart of the Rescorla-Wagner (1972) model. In terms of Wagner's (1978, 1981) development

of this model, it is necessary for information about the occurrence of the signal and reinforcer to be processed conjointly if an increment in the associative knowledge about their relationship is to result from a pairing. If either the signal or the reinforcer fails to receive adequate processing, the animal will learn little about their association. The major factor determining whether a signal or reinforcer will be processed adequately is its surprisingness, for only unexpected stimuli will receive sufficient processing to sustain substantial learning.

As we have already seen, the effect of varying the surprisingness of the reinforcer is clearly illustrated by the blocking effect. On the initial AB compound trials of the second stage (see Table 3), the target or added signal A will be adequately processed by the animals in all conditions, for the occurrence of this stimulus is equally unexpected for all animals. By contrast, the reinforcer is differentially predicted in the various groups. Like signal A, the reinforcer is unexpected for the control animals and, therefore, should receive adequate processing. For the blocking group, however, the occurrence of the reinforcer is well predicted by the pretrained signal B as a result of the learning that took place in the first stage and, thus, the reinforcer will receive insufficient processing to sustain learning about its relationship with the added signal A during the second stage.

Blocking represents an effect that, according to this theory, is predominantly due to variations in the processing or effectiveness of the reinforcer. The consequences of varying the processing of the signal, on the other hand, are to be found in the phenomenon of latent inhibition. The fact that preexposure to a number of presentations of a signal alone retards subsequent learning involving that stimulus must be due to, in some way or another, a change in the effectiveness of that stimulus. This idea is often expressed by saying that latent inhibition reflects a loss in the *associability* of the signal, as though there was attached to each stimulus a parameter that determines how readily it can be associated with the reinforcer. According to Wagner, this loss occurs because during the preexposure stage the animals learn to expect the signal within the particular experimental context employed, so that when it is now paired with the reinforcer, its occurrence is no longer surprising. In turn, this retards processing the stimulus and hence learning about its association with the reinforcer.

Blocking and latent inhibition represent but two of a plethora of conditioning effects that can be encompassed by Wagner's development of the Rescorla-Wagner theory, and there is little doubt that at present this account is the most influential within the area of conditioning.

THE PEARCE-HALL THEORY

Wagner's account argues that variations in conditioning are due to variations in the processing of the signal and the reinforcer. Pearce and Hall (1980) asked, however, whether it is really necessary to assume that the effectiveness of both stimuli can be changed as a result of experience. They noted that in latent inhibition and in blocking the target signal is presented in a context where nothing unexpected occurs; in latent inhibition training no other events are presented, whereas the target or added signal A in the blocking procedure is presented in association with two

other stimuli, the pretrained signal B and the reinforcer, both of which are expected on the basis of the animal's past experience. Perhaps, Pearce and Hall argued, a signal progressively loses its ability to engage the learning process and hence suffers a loss of associability whenever it is presented in a context where nothing surprising happens.

Latent inhibition follows directly from such an assumption. Across the series of presentations of the signal alone during preexposure, the signal should lose its associability so that when it is paired with the reinforcer for the first time, it will no longer be processed in a manner that allows the animal to learn about its new association with the reinforcer. Of course, the occurrence of the reinforcer will be surprising, so that the associability and processing of the signal should be restored, but by that time the rate of conditioning will have been retarded. In latent inhibition the loss in associability occurs before any conditioning to the signal. By contrast, in blocking the decrement in the processing of the target stimulus A actually occurs during conditioning to this stimulus in the second, compound conditioning stage of the experiment. The only difference between the blocking and control conditions is that this loss occurs more rapidly in the former because all the events occurring in association with signal A are fully expected from the outset of compound training in the blocking group. This enhanced loss of associability in turn restricts the amount of conditioning to the signal A in this group.

From this analysis it can be seen that Pearce and Hall make what at first sight appears to be a very surprising claim. Even during simple conditioning there should be a progressive loss in the processing of the signal as the animal learns about the relationship between the signal and reinforcer. This is because nothing surprising occurs in association with the signal when it is well established as a predictor of the reinforcer. Of course, the animal continues to respond to the signal, but Pearce and Hall argue that this represents just an automatic, although conditional, response to the signal. Moreover, they provide some independent evidence for such a decline by measuring the *orienting* response to the signal. Many stimuli elicit responses that seem to be related to their processing and hence their associability; for instance, when a localized light is first presented to a rat, it looks at the stimulus. If Pearce and Hall are right in supposing that signal processing is reduced during simple conditioning, we should expect to see a progressive decline in the orienting response as the conditional response develops. This is just what Kaye and Pearce (1984) found when they made a light consistently signal the delivery of food. Moreover, they also demonstrated that by associating the signal with an inconsistent outcome the orienting response could be maintained. This they did by reinforcing only half of the signal presentations so that it was always paired with a surprising or unexpected outcome. In contrast to consistent reinforcement, this partial schedule maintained a high and constant level of orientation to the light.

Thus, Pearce and Hall would argue that many of the critical phenomena of conditioning, such as latent inhibition and blocking, can be explained simply in terms of variations in the processing of the signal, leaving the changes in reinforcer effectiveness posited by Wagner's theory somewhat redundant. However, there are important phenomena, learned irrelevance and causal relevance for example, that lie outside

the scope of both of the accounts we have considered so far. The main attempt to address these effects has been made by Mackintosh (1973, 1975).

THE MACKINTOSH THEORY

There is something very counterintuitive about the claim that a stimulus of no predictive significance receives the same treatment by the learning mechanism as a signal for a biologically important event. And yet this is the assumption that Pearce and Hall make when they argue that after conditioning and latent inhibition stimuli have a low associability and are resistant to entering into association with a new reinforcer. Although this claim seems reasonable for a stimulus of no predictive significance (i.e., a latently inhibited stimulus), one would expect an animal to learn readily about relationships involving a stimulus that in the past has signaled important events. This is the position taken by Mackintosh (1973, 1975).

Basically, Mackintosh argued that the associability of a stimulus is determined by how good a predictor it is relative to the other stimuli present in a conditioning episode. If it is the best predictor its associability goes up, whereas if it is worse (or even no better) than the other stimuli present, its associability suffers a decline. As with the Pearce-Hall theory, blocking is due to a decline in the associability of the added stimulus A because it is a worse predictor of the reinforcer than the pretrained stimulus B during compound conditioning in Stage 2 (see Table 3). In contrast to the Pearce-Hall account, however, simple conditioning leads to an increment in the associability of the signal for, in this case, it is the best predictor of the reinforcer.

The fact that Mackintosh argued that the associability of a stimulus records its predictive history allows this theory to make contact with phenomena that appear to reflect this history. For instance, the observation that animals learn more readily about an association between a flavor and illness than one between the same taste and an exteroceptive reinforcer must, in some way or another, reflect the predictive history of these signals during either the ontogeny of the individual animal or the phylogeny of the species. To encompass this causal relevance effect, we must allow each potential signal to possess a number of associability values, one for each class of reinforcers (see Dickinson & Mackintosh, 1979). The fact that gustatory stimuli and gastric consequences will have been highly correlated during development should increase the associability of this class of signal for this type of reinforcer. Similarly, the independence of exteroceptive stimuli in respect of tastes should ensure that the corresponding associabilities decline as is observed in the learned irrelevance effect. It is clear, however, that causal relevance is unlikely to be entirely a matter of experience; one-day-old rats appear to be predisposed to make causally relevant associations (Gemberling & Domjan, 1982). In this case, it appears that information about variations in stimulus-reinforcer relationships that has been gained across phylogeny can be transmitted by the inheritance of the appropriate starting values for associability.

Despite the major differences between the theories I have outlined, they have in common one important assumption; conditioning involves two distinct forms of learning. As well as learning about the relationship between a signal and a particular

reinforcer, the way in which the animal processes the signal itself changes. Such an assumption appears to be essential if a theory is to get grips with the major discovery in conditioning during the last 20 years, namely that conditioning to a stimulus is selective in accord with its predictive history. Where the theories differ is in the rules governing these changes in processing or associability. To recapitulate, for Wagner the critical factor is whether or not a stimulus itself is predictable; only unexpected stimuli have a high associability. By contrast, Mackintosh emphasizes predictive power with a high associability being assigned to a stimulus having a history of being a good predictor of the reinforcer. Pearce and Hall also claim that associability depends on the history of the signal, but in this case the important factor is whether or not the stimulus has been presented in conjunction with an unexpected reinforcer in the recent past. At present, it is impossible to make any final adjudication between these accounts. No one theory has, as yet, received universal acceptance and all have their particular strength and weakness within the various domains of conditioning.

INHIBITORY CONDITIONING

As we have seen, one of the major discoveries of the last 20 years is that conditioning is sensitive to the correlation between a signal and reinforcer, with positive correlations leading to conditioning relative to an uncorrelated or noncontingent schedule. This sensitivity to event correlations immediately raises the question of what would happen if we exposed an animal to a negative correlation. Such a relationship is represented by the bottom sequence in Figure 1. Here the reinforcer occurs in the absence of the signal A, but never in its presence, so that the stimulus signals a reduction in the likelihood of the reinforcer.

We have already seen in Rescorla's (1971) superconditioning experiment that animals appear to be capable of learning about negative correlations. In fact, Rescorla (1969) was the first to investigate systematically learning about such relationships. He found that an auditory signal that had been trained under a negative correlation with the reinforcer had little or no behavioral effect when it was presented by itself outside the training context. From a cognitive point of view this is unsurprising. If the animal learns that the stimulus signals the nonoccurrence of the reinforcer, we should expect such knowledge to control behavior only in a context in which it is relevant. Such a context, of course, is one in which the animal has reason to expect the reinforcer to occur in the first place; in the absence of such an expectation, a signal predicting reinforcer omission is of little significance. So to test his rats' knowledge about the negative correlation, Rescorla established a second signal, a light, as a predictor of the reinforcer by pairing them, and then compared the magnitude of the conditional response to the light alone with that elicited by a compound of the light and the auditory stimulus that was originally trained under the negative correlation. To the extent that the animals had learned about this correlation, the expectation maintained by the tone should have negated that controlled by the light, thereby inhibiting the response to the light. This is just what Rescorla observed with the magnitude of the inhibitory effect being systematically related to the degree of the negative correlation during training. Signals, such as the tone, that acquire this

inhibitory property through an associative relationship with the reinforcer are referred to as *conditional inhibitors* to distinguish them from *conditional excitors*, such as the light, which elicit the target conditional response.

As in the case of positive event correlations, there is no need to assume that inhibitory conditioning reflects a direct sensitivity to negative relationships. Figure 1 makes clear that a negative correlation is realized by presenting the animals with two types of intermixed episodes, one in which the contextual stimuli, represented by signal B, are paired with the reinforcer and one in which a compound of stimuli A and B is nonreinforced. We can represent such a sequence as a B+, AB- schedule. Once the negative correlation is analyzed in these terms, we can see that the primary condition for establishing inhibitory conditioning is the simple antithesis of that for the excitatory variety. Whereas a conditional excitor results from pairing it with the unexpected presentation of a reinforcer, inhibitory conditioning is established by pairing the target with the surprising omission of the reinforcer. The B+ episodes lead the animal to expect the reinforcer in the presence of stimulus B so that its omission following the AB compound is unexpected. If this analysis is correct, we should be able to establish stimulus A as a conditional inhibitor by using a discrete stimulus as signal B rather than the background cues. In fact, the effectiveness of this B+, AB- schedule has long been known for it represents Pavlov's classic conditioned inhibition procedure.

The three theories (Mackintosh, 1975; Pearce & Hall, 1980; Wagner, 1981) that we considered in the previous section all endorse this analysis of inhibitory conditioning. Moreover, to the extent to which they address this issue at all, they assume that the associative knowledge about negative correlations is represented independently of that about positive relationships. This means that exposure to a negative correlation sets up an associative representation of this relationship that is independent of any knowledge the animal might have about a positive relationship between the signal and reinforcer. These two representations only interact to control the performance of a conditional response. Such theories have profound implications for the nature of extinction. When a previously established excitor is no longer reinforced, the conditions for the development of inhibition are fulfilled; a reinforcer is omitted at the time when the animal expects it. As a result, the extinction of the conditional response is due to the acquisition of inhibition that counteracts the original excitatory properties of the signal. And in this sense, none of the theories allow for unlearning.

OCCASION SETTING

Although this analysis of inhibitory conditioning commands general acceptance, there is no doubt that it is oversimplified. This point is clearly illustrated in a recent study by Holland (1984). He trained rats on the standard conditional inhibition schedule in which reinforced presentations of the excitor, B+, were intermixed with nonreinforced presentations of a simultaneous compound of the excitor and inhibitor, AB-. Not surprisingly, the animals learned to discriminate these two types of trials, responding to presentations of B alone, but not to the AB compound. When he

subsequently reinforced stimulus A by itself, he found that acquisition of excitatory conditioning was retarded. This retardation effect follows from the fact that stimulus A had become a conditional inhibitor during discrimination training; if the animals initially thought that stimulus A signalled the omission of the reinforcer, it would require additional reinforced trials in order to build up sufficient excitatory strength to overcome the initial bias. But once this had been done, stimulus A should have become a net excitor and so have lost its inhibitory capacity. As predicted, Holland found that signal A was no longer capable of inhibiting responding when presented in compound with the original excitor.

So far, Holland's results fit the standard analysis perfectly. The problem comes with his second group of rats. These were exposed to exactly the same discrimination schedule as the first except that stimuli A and B were presented in a serial rather than simultaneous compound on nonreinforced trials with stimulus A occurring before stimulus B. Now when he reinforced stimulus A alone, there was no retardation of excitatory conditioning. Moreover, representing the two stimuli in serial compound (and, for that matter, in simultaneous compound) revealed that stimulus A was still fully capable of inhibiting the excitatory responding controlled by stimulus B even though it itself was now a strong excitor for the same response. It is as though the animals had acquired two independent pieces of knowledge about stimulus A. On the one hand, they knew that this signal predicted the reinforcer and, on the other, that it also predicted that stimulus B would not be reinforced. Furthermore, it appears from Holland's results that each of these items of knowledge can be changed independently of the other.

In summary, two different forms of learning occur, depending on whether a simultaneous or serial AB compound is used during training. During training with a simultaneous compound, the animal learns that stimulus A signals that an expected reinforcer will not occur so that this stimulus can inhibit the action of any conditional excitor for the same reinforcer. Furthermore, converting stimulus A itself into an excitor counteracts its inhibitory property. By contrast, following serial training stimulus A signals a relationship between a particular excitor and the reinforcer rather than the simple omission of the reinforcer. As a result, its inhibitory property will not act on another simple excitor nor will it be altered by changing its predictive status with respect to the reinforcer itself. Holland refers to this type of conditional control as "occasion setting"; stimulus A sets the occasion when stimulus B will not be reinforced.

It is now clear that occasion setting is not restricted to the negative relationship studied by Holland (1984). Rescorla (1985), for instance, has reported positive occasion setting by using an AB+, B- schedule with pigeons in which stimulus A signals that stimulus B will be reinforced. As in the negative case, the control exercised by stimulus A over responding to stimulus B is independent of the direct relationship between signal A and the reinforcer. Independent reinforcement and nonreinforcement of stimulus A by itself appears to have no effect on the ability of this signal to control reactions to stimulus B.

Besides its general functional importance, occasion setting is significant because it lies outside the scope of our current theoretical analyses of conditioning. Up to

now, we have thought of the knowledge underlying conditioning as a two-term associative structure between the signal and the reinforcer, and yet occasion setting appears to involve three terms: the occasion setter, the signal, and the reinforcer. One way to harmonize occasion setting with the standard analysis is to assume that the compound presentation of the occasion setter A and the signal B forms a single configural cue, so that the AB+ B- schedule effectively presents the animal with a discrimination between two stimuli, AB and B. If this is so, we should not observe transfer between occasion setters because a compound of the signal and the occasion setter would not form a reinforced configuration unless they had been trained together. To test whether such transfer occurred, Rescorla (1985) trained his pigeons on two conditional discriminations concurrently, AB+ B- and CD+ D-, in which the occasion setters A and C were dissimilar stimuli from different modalities. After the birds had learned these discriminations, he presented them with AD and CB compounds for the first time. According to the configural theory, these compounds, being novel and without a history of reinforcement, should control little responding, whereas, in fact, Rescorla found almost as much responding to these novel compounds as to the original ones. Note that this transfer is not at variance with Holland's failure to find that an occasion setter would fail to exert control over a simple excitator with which it had not been trained. Control appears to transfer to other excitators as long as they themselves have been trained within an occasion setting relationship.

These findings clearly point to a new form of associative learning in conditioning involving three rather than two terms which allows the animal to use one stimulus, the occasion setter, to predict when another will be reinforced (or nonreinforced). Beyond this, we can say little at present for we have no idea about the structure of the associative knowledge underlying occasion setting nor about the processes mediating the acquisition of this knowledge.

INSTRUMENTAL CONDITIONING

It is notable that the discussion so far has been restricted to classical or Pavlovian conditioning in which the subjects simply experience a relationship between a stimulus and a reinforcer, neither of which is under their control. Except for the initial discussion of reinforcer devaluation, nothing has been said about instrumental or operant conditioning in which behavior changes as a result of a contingency between one of the subject's own actions and the reinforcer. The reason for this neglect is simple; the contemporary study of the associative learning underlying conditioning has been conducted almost exclusively within the Pavlovian paradigm. This emphasis has been due largely to the technical reason that only in the Pavlovian procedure does the experimenter have full control over the events and episodes experienced by the subject and is therefore the preferable procedure for studying associative learning. But it does leave open the question of the extent to which common principles of learning underly the two forms of conditioning.

It has often been suggested that Pavlovian and instrumental conditioning, although operationally distinct, do in fact represent the same process. Thus, for

example, Pavlov's dogs may have salivated to a signal for food because this response, in some way or another, enhanced the value of the food or their ability to cope with this reinforcer. According to this analysis, the conditional response is not controlled directly by the Pavlovian signal–reinforcer relationship, but rather by an instrumental contingency between anticipatory salivation and the enhancement of the value of the reinforcer. However, the fact that responding persists under what is called an omission schedule shows that this is not so. Even if we ensure that every time the dog salivates in response to the signal the food is omitted, so that anticipatory salivation can never enhance the value of the food, the animal continues to salivate to the signal (Sheffield, 1965). So, it appears that conditional responding can be controlled in certain cases directly by the stimulus–reinforcer relationship, whereas in others it is the action–reinforcer contingency that is important.

The fact that instrumental and Pavlovian conditioning can be distinguished in terms of the controlling relationship does not imply, however, that the underlying learning processes necessarily differ. In fact, what evidence we have suggests that they are similar. We have seen that Pavlovian conditioning is sensitive to the correlation between the signal and reinforcer when the probability of contiguous signal–reinforcer pairings is kept constant. In an analogous manner, instrumental conditioning is impaired when the action–reinforcer correlation is degraded without altering the probability of reinforcement paired with the action (e.g., Dickinson & Charnock, 1985; Hammond, 1980). Moreover, similar principles of selective learning appear to operate in the two cases. Just as the presence of an alternative signal for the reinforcer can reduce Pavlovian conditioning to a target stimulus, such a signal can also compete with an instrumental action to impair instrumental conditioning (Mackintosh & Dickinson, 1979; Pearce & Hall, 1978).

The instrumental analogue of learned irrelevance can be found in the well-known learned helplessness effect (Maier & Seligman, 1976). Exposing an animal to a zero correlation between its own activity and a reinforcer, usually an aversive one, retards subsequent learning about a relationship between a particular action and the reinforcer. Furthermore, a parallel to the causal relevance effect might be seen in Shettleworth's (1975, 1978) demonstrations that various behavior patterns are differentially sensitive to instrumental reinforcement in a way that depends, at least in part, upon the type of reinforcer employed.

Instrumental avoidance schedules arrange a negative correlation between an animal's action and the reinforcer in a manner that parallels Pavlovian inhibitory conditioning. Just as the conditional inhibitor predicts the omission of the reinforcer in the Pavlovian case, so the instrumental action causes the omission of the reinforcer under the avoidance schedule. This parallel is clearly demonstrated by the fact that a feedback stimulus produced by the avoidance response acquires Pavlovian inhibitory properties (Weisman & Litner, 1972). Finally, the phenomenon of occasion setting in Pavlovian conditioning can be identified with that of discriminative control in the instrumental variety. A discriminative stimulus is a signal that "sets the occasion" on which an instrumental action is reinforced (or nonreinforced).

Given these obvious parallels between the phenomena of the two forms of conditioning, there are good grounds for arguing that they are mediated by a common form of associative learning. All that differs is the elements of association; in the

Pavlovian case the subject learns about the association between a signal and reinforcer and possibly the relationship between this association and an occasion setter, whereas the learned association is between an action and the reinforcer in the instrumental case which, in turn, may be related to a discriminative stimulus.

CONCLUSIONS

I introduced this brief survey of contemporary views of conditioning with reference to the question of whether conditioning could still be regarded as an important process in the etiology of behavioral disorders and the effectiveness of therapeutic procedures. The discussion has, I hope, brought out two major points that are relevant to this question. The first is that even in its paradigmatic cases conditioning cannot be understood without reference to cognitive processes. Secondly, conditioning is neither a simple nor well-understood process; the last 20 years have revealed complexities and subtleties that transcend any simple reinforcement mechanism. Both these points deserve some comment.

At one time, conditioning appeared to be defined as an unconscious, noncognitive, and automatic process. So prevalent was this opinion that no more than 12 years ago Brewer (1974) could claim that "There is no convincing evidence for operant or classical conditioning in adult humans." This claim was based, not on the failure to observe the appropriate behavioral changes in humans, but on the fact that all such changes appeared to be mediated by cognitions. So strong was the stimulus-response theory of conditioning in the intellectual heritage of American psychology that Brewer failed to distinguish the empirical phenomenon of conditioning from the reinforcement theory of the time. As we have seen, however, the contemporary approach recognizes a role for representations and beliefs about predictive and causal relationships even in the case of animal conditioning.

This is important because it means that the conditioning model is not necessarily at variance with the increasing recognition of the role of mental processes in psychiatric disorders that is reflected in the growth of cognitive therapy (Gelder, 1985). The contemporary model allows for the idea that inappropriate and maladaptive beliefs may well have arisen through a conditioning experience and that such beliefs may be changed through conditioning procedures. Nor should the model be taken as endorsing only behavioral techniques in therapy; there is no reason why alterations in conditional behavior should not follow cognitive readjustments.

It must be recognized, however, that although contemporary theory provides a more liberal model of conditioning than the traditional one, it is also a more complex and uncertain one. Whereas the behavior therapists of the previous generation could treat conditioning as a simple and well-understood phenomenon, they will now find in the literature a plethora of effects competing for a variety of theories; a stimulus can be a conditional excitor, a conditional inhibitor, or even an occasion setter; it can have a high associative strength but low associability or any other combination of these properties. Whether or not this richer theoretical framework can provide a rational basis for the development of more effective therapeutic procedures remains an open question; at the very least, the contemporary approach stands a better chance

of matching up to the complex cognitive and behavioral profiles observed in the clinic than did traditional reinforcement theory.

REFERENCES

- Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Quarterly Journal of Experimental Psychology*, *33B*, 109–122.
- Brewer, W. F. (1974). There is no convincing evidence for operant or classical conditioning in adult humans. In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes*, (pp. 1–42). Hillsdale, NJ: Erlbaum.
- Colwill, R. C., & Rescorla, R. A. (1985). Postconditioning devaluation of a reinforcer affects instrumental responding. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 120–132.
- Dickinson, A. (1980). *Contemporary animal learning theory*. Cambridge: Cambridge University Press.
- Dickinson, A. (1985). Actions and habits: The development of behavioural autonomy. In L. Weiskrantz (Ed.), *Animal intelligence*, Oxford: Clarendon Press.
- Dickinson, A., & Charnock, D. J. (1985). Contingency effects with maintained instrumental reinforcement. *Quarterly Journal of Experimental Psychology*, *37B*, 397–416.
- Dickinson, A., & Mackintosh, N. J. (1979). Reinforcer specificity in the enhancement of conditioning by posttrial surprise. *Journal of Experimental Psychology: Animal Behavior Processes*, *5*, 162–177.
- Domjan, M., & Wilson, N. E. (1972). Specificity of cue to consequence in aversion learning in the rat. *Psychonomic Science*, *26*, 83–95.
- Gelder, M. (1985). Cognitive therapy. In K. Granville-Grossman (Ed.), *Recent advances in clinical psychiatry* *5*, (pp. 1–21). Edinburgh: Churchill Livingstone.
- Gemberling, G. A., & Domjan, M. (1982). Aversion learning in one-day-old rats: Taste-toxicosis and texture-shock associations. *Journal of Comparative and Physiological Psychology*, *96*, 105–113.
- Hammond, L. J. (1980). The effect of contingency upon appetitive conditioning of free-operant behavior. *Journal of the Experimental Analysis of Behavior*, *34*, 297–304.
- Hammond, L. J., & Paynter, W. E. Jr. (1983). Probabilistic contingency theories of animal conditioning: A critical analysis. *Learning and Motivation*, *14*, 527–550.
- Holland, P. C. (1981). Acquisition of representation-mediated conditioned food aversions. *Learning and Motivation*, *12*, 1–18.
- Holland, P. C. (1984). Differential effects of reinforcement of an inhibitory feature after serial and simultaneous feature negative discrimination training. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 461–475.
- Kamin, L. J. (1969). Predictability, surprise, attention and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279–296). New York: Appleton-Century-Crofts.
- Kaye, H., & Pearce, J. M. (1984). The strength of the orientating response during Pavlovian conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 90–109.
- Kimble, G. A. (1961). *Hilgard and Marguis' conditioning and learning*. New York: Appleton-Century-Crofts.
- Lubow, R. E. (1973). Latent inhibition. *Psychological Bulletin*, *79*, 398–407.
- Mackintosh, N. J. (1973). Stimulus selection: Learning to ignore stimuli that predict no change in reinforcement. In R. A. Hinde & J. Stevenson-Hinde (Eds.), *Constraints on learning* London: Academic Press.
- Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*, 276–298.
- Mackintosh, N. J. (1983). *Conditioning and associative learning*. New York: Oxford University Press.
- Mackintosh, N. J., & Dickinson, A. (1979). Instrumental (Type II) conditioning. In A. Dickinson & R. A. Boakes (Eds.), *Mechanisms of learning and motivation*, Hillsdale, NJ: Erlbaum.
- Maier, S. F., & Seligman, M. E. P. (1976). Learned helplessness: Theory and evidence. *Journal of Experimental Psychology: General*, *105*, 3–46.
- Pearce, J. M., & Hall, G. (1978). Overshadowing the instrumental conditioning of a lever press response by a more valid predictor of reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, *4*, 356–367.

- Pearce, J. M. & Hall, G. (1980). A model of Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*, 532–552.
- Rescorla, R. A. (1968). Probability of shock in the presence and absence of the CS in fear conditioning. *Journal of Comparative and Physiological Psychology*, *66*, 1–5.
- Rescorla, R. A. (1969). Conditioned inhibition of fear resulting from negative CS-US contingencies. *Journal of Comparative and Physiological Psychology*, *67*, 504–509.
- Rescorla, R. A. (1971). Variations in the effectiveness of reinforcement and nonreinforcement following prior inhibitory training. *Learning and Motivation*, *2*, 113–123.
- Rescorla, R. A. (1980). *Pavlovian second-order conditioning: Studies in associative learning*. Hillsdale, NJ: Erlbaum.
- Rescorla, R. A. (1985). Conditioned inhibition and facilitation. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* Hillsdale, NJ: Erlbaum.
- Rescorla, R. A., & Solomon, R. L. (1967). Two-process learning theory: Relationship between Pavlovian conditioning and instrumental learning. *Psychological Review*, *74*, 151–182.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.
- Revusky, S.H. (1971). The role of interference in association over a delay. In W. K. Honig & P. H. R. James (Eds.), *Animal memory* (pp. 155–213). New York: Academic Press.
- Revusky, S. H., & Garcia, J. (1970). Learned associations over long delays. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 4, pp. 1–84). New York: Academic Press.
- Sheffield, F. D. (1965). Relation between classical conditioning and instrumental learning. In W. F. Prokasy (Ed.), *Classical conditioning: A symposium* (pp. 302–322). New York: Appleton-Century-Croft.
- Shettleworth, S. J. (1975). Reinforcement and the organization of behavior in golden hamsters: Hunger, environment, and food reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, *1*, 56–87.
- Shettleworth, S. J. (1978). Reinforcement and the organization of behavior in golden hamsters: Punishment of three action patterns. *Learning and Motivation*, *9*, 99–123.
- Testa, T. J. (1974). Causal relationships and the acquisition of avoidance responses. *Psychological Review*, *81*, 491–505.
- Tolman, E. C. (1933). Sign-gestalt or conditioned reflex? *Psychological Review*, *40*, 246–255.
- Trapold, M. A., & Overmier, J. B. (1972). The second learning process in instrumental learning. In A. H. Black & W. K. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 302–322). New York: Appleton-Century-Crofts.
- Wagner, A. R. (1978). Expectancies and the priming of STM. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior* (pp. 177–209). Hillsdale, NJ: Erlbaum.
- Wagner, R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms*, Hillsdale, NJ: Erlbaum.
- Weisman, R. G., & Litner, J. S. (1972). The role of Pavlovian events in avoidance training. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning*, (pp. 253–270). New York: Academic Press.

CHAPTER 4

A Primate Model of Phobic Fears

Susan Mineka

INTRODUCTION

The use of animal models in clinical psychology and psychiatry has a long and mixed history. There were early proposals by Watson and Rayner (1920) and by Pavlov (1927) that classical conditioning plays a prominent role in the origins of a variety of so-called neurotic disturbances. These were soon followed by numerous demonstrations of “experimental neurosis” in which disturbed behavior was induced in a range of different species through exposure to a wide range of somewhat aberrant conditioning procedures. Research on this topic was performed in a number of well-known laboratories, including those of Pavlov, Liddell, Masserman, N. R. F. Maier, and Wolpe (see Broadhurst, 1960, 1973; Mineka & Kihlstrom, 1978, for reviews).

This early work on experimental neurosis constituted the first attempt at developing animal models of human psychopathology. And indeed it was quite influential in helping to establish the foundations of behavioral approaches to understanding the etiology of neurotic disorders such as anxiety states and depression. Related work on avoidance and on extinction of fear also helped to establish the foundations of behavioral approaches to the treatment of such disorders (e.g., Baum, 1970; Solomon, Kamin, & Wynne, 1953; Wolpe, 1958). Unfortunately, the early work on experimental neurosis, and therefore on animal models of psychopathology *per se*, was fairly unsystematic and consequently fell into disfavor for a number of years. This resulted primarily from the failure of these investigators to document whether compelling phenotypic or functional similarities existed between the animal’s disturbed behavior and the supposedly parallel human disorders. (See Mineka, 1982, 1985; Mineka & Kihlstrom, 1978; for further discussions of this early history of the use of animal models.)

SUSAN MINEKA • Department of Psychology, University of Texas, Austin, Texas 78712. Preparation of this chapter was supported by Grant BNS-8507340 from the National Science Foundation to the author. The research described in this chapter has been generously supported by grants from the University of Wisconsin Graduate School and by Grants BNS-7823612, BNS-8216141, and BNS-8507340 from the National Science Foundation to the author.

Since 1970 the use of animal models of psychopathology has regained some popularity as several investigators have more systematically explored causal factors involved in the development of abnormal behavior in animals. Concurrently they have also attempted to document more compelling similarities to the parallel human disorders. For example, McKinney (1974) and Seligman (1974, 1975) made compelling arguments that animal models can be useful if certain criteria are adhered to in the development of the model. In particular, they proposed that in developing an animal model one should attempt to document similarities in the symptomatology, the etiology, the prevention, and the therapy for the disorder. Obviously not all of the parallels will be possible to detail at the outset, because much may be unknown about some of these factors (e.g., prevention) for either the animal or the human disorder. However, herein lies one of the special advantages of developing an animal model. Initially some compelling similarities must be drawn between the animal model and the human disorder using several of these criteria. However, the animal model can then be used to test hypotheses about other possible parallels (e.g., prevention) that cannot feasibly be tested experimentally with humans.

Animal research on fear conditioning and extinction has had a large impact on the development of behavioral models of the origins and therapy for fear and anxiety disorders. This has occurred in spite of the fact that very little of this research has used the criterion approach advocated by Seligman and McKinney. (See Mineka, 1985, for a fairly comprehensive review.) Instead, most of this research has involved the development of what have been called minimodels (Marks, 1977; Mineka, 1985) that help to illuminate certain aspects of the symptomatology, or the etiology, or the therapy for these disorders. Minimodels are simply conditioning phenomena that cannot by themselves account for all aspects of the etiology of the disorder, or the full range of symptomatology of the disorder, or the complete mechanisms through which therapeutic benefits are produced. Instead, they help to illuminate some of the most prominent features of the origins or treatment of the disorder. Ideally these conditioning phenomena should each be thought of as one step in a complex sequence or interaction of events that may be involved in the etiology, maintenance, or therapy for a disorder.

One of the primary reasons researchers have been restricted to the use of minimodels stems from the fact that a great majority of such research has used rats as subjects (or occasionally dogs or cats). There are inherent limitations in the number of compelling similarities that can be drawn on these four criteria when the species are only so remotely related to one another. We have been fortunate to have access to rhesus monkeys who share many more behaviors and features of social development with humans than do rats, dogs, or cats. Consequently in our research we have not been restricted to the use of minimodels. Indeed we have developed a primate model of phobic fears that draws parallels to human simple phobias on all four of the criteria discussed earlier: symptomatology, etiology, therapy, and prevention. Our model does not account for every aspect of the symptomatology, or every etiological pathway, or every possible therapy, or every possible form of prevention for simple phobias in humans. It does, however, provide a compelling model of many of the cardinal symptoms of phobias and of how many phobias may originate. It also provides a

model for how therapy may often operate to produce its beneficial effects, and for how the acquisition of many phobias could be prevented.

The specific fear—snake fear—that we chose to use in the development of our primate model of phobic fears is one that has received considerable attention in the primate literature over the past 50 years. Controversy on this topic has largely centered around the issue of whether the fear of snakes exhibited by many primate species is innate or learned. The most recent evidence on this issue strongly suggests, at least for rhesus and squirrel monkeys, that this fear is learned. This conclusion stems from observations that only monkeys reared in the wild exhibit a pronounced fear of snakes. The failure of laboratory-reared monkeys to exhibit a fear of snakes could be an aberration from normal development. However, the more likely explanation of this pattern of findings is that the lab-reared monkeys simply lacked the necessary learning experience that the wild-reared monkeys had had to acquire the fear (cf. Joslin, Fletcher, & Emlen, 1964; Mineka, Keir, & Price, 1980; Murray & King, 1973).

The advantage of using snake fear for our primate model was that we had access both to monkeys that already exhibited an intense fear of snakes (wild-reared monkeys) and to monkeys that did not exhibit any fear of snakes (lab-reared monkeys). This gave us the possibility to study therapy in the wild-reared monkeys, and to study etiology and prevention in lab-reared monkeys. It also gave us the opportunity to study an etiological pathway about which relatively little was known, namely acquisition of a fear through observational conditioning. This latter goal could be accomplished through the combined use of wild-reared monkeys as models exhibiting their fear of snakes, and of lab-reared monkeys as observers who could watch the wild-reared models behaving fearfully with snakes.

SYMPTOMATOLOGY OF PRIMATE FEAR OF SNAKES

PARALLELS WITH HUMAN PHOBIAS USING LANG'S THREE-SYSTEMS MODEL OF FEAR

For some years researchers and clinicians have known that a client's self-report about the intensity of his or her fears does not always accurately reflect the real level of stress or interference the fear or phobia is creating in the person's everyday life. This example illustrates that fear is not a hard phenomenal lump that can be directly accessed through self-report, but rather a set of loosely coupled response components (cognitive/subjective, physiological, and behavioral/avoidant). Such a view has been carefully described by Lang and his colleagues (1968, 1971, 1985). Lang's research, as well as that of numerous other researchers (e.g., Hodgson & Rachman, 1974; Mineka, 1979; Rachman & Hodgson, 1974) has clearly shown that these fear response systems do not always covary together. A client may, for example, report high levels of subjective distress but show little difficulty in approaching his feared object and/or show few signs of physiological arousal in the presence of that object, or vice versa.

Behavior therapy researchers have also amply documented that treatments designed to treat a client's fear or phobia often produce improvements primarily in one fear-response system, with improvement in the other fear-response systems lagging considerably behind. Flooding or exposure therapies, for example, often produce their first effects on reducing behavioral avoidance of the feared object. This may leave the client with significant levels of subjective distress and/or physiological arousal unless treatment is continued until fear in these response systems is also reduced (e.g., Hodgson & Rachman, 1974; Rachman, 1978).

Important implications of this work for the present chapter are two-fold. First, it is clear that primate research designed to test hypotheses about acquisition of phobic fears in humans should document that the fear being modeled exists in more than one of the three fear-response systems. Second, attempts to model therapy and prevention processes in monkeys must monitor fear in several response systems to assure that the fear has been successfully treated or prevented. Therefore, in all of the research discussed in the following review, we have monitored fear in two of the fear-response systems—behavioral avoidance and behavioral distress.

Measurement of Fear in Monkeys. In our research at Wisconsin's Harlow Primate Laboratory, behavioral avoidance is measured in two different contexts. First, in the Sackett Circus apparatus the monkeys are placed in a central compartment that is surrounded by four outer compartments. On the outside of the far wall (made of Plexiglas) of these outer compartments are placed a variety of neutral and snake stimuli (see Mineka, Davidson, Cook, & Keir, 1984, for details). The monkeys are given 5 minutes in which they can wander freely between the central compartment and the four outer compartments. When they are not afraid of any of the objects placed outside the Plexiglas walls, they tend on the average to spend equal amounts of time in the four outer compartments. By contrast, when they are afraid of one or more of the stimuli placed outside the Plexiglas walls, they tend to escape from or completely avoid those compartments, and spend the great majority of their time in the compartments with neutral stimuli. Thus, in the Sackett Circus behavioral avoidance of snakes is indexed by small amounts of time spent in the snake compartments.

The second context in which we measure behavioral avoidance is the Wisconsin General Test Apparatus (WGTA; Harlow, 1949). In this apparatus monkeys are pretrained to reach rapidly across an open Plexiglas box to obtain a desired food treat (see Mineka et al., 1980, 1984, for details). When a neutral object that elicits no fear is placed in the open Plexiglas box, the monkeys generally reach for the food treat within a few seconds. By contrast, when a feared object (such as a snake) is placed in the open Plexiglas box the monkeys show great reluctance to reach for the food, and generally do not respond within the 60 second maximum duration of a trial. Thus, the conflict created by fear of the snake and desire to approach the food results in behavioral avoidance, indexed by long food-reach latencies.

In the WGTA we also have a measure of behavioral disturbance or distress, probably measuring what Lang (1968, 1971) describes in humans as the cognitive/subjective fear-response system. While the monkeys' are being monitored for their food reach latency in the presence of feared and neutral objects, the experimenter also carefully observes and records a dozen different fear or disturbance behaviors that have been shown to occur when monkeys are confined within a few feet of a

feared object. These fear or disturbance behaviors include fear grimacing, threat faces, clutching the cage, staring, eye aversion, sudden retreat to the back of the cage, lip smacking, and piloerection (see Mineka *et al.*, 1980; Mineka *et al.*, 1984, for details). Typically, high levels of these disturbance behaviors occur in the presence of feared objects and few or none occur in the presence of nonfeared neutral objects.

We have not directly monitored physiological arousal in order to tap the physiological fear response system. However, it seems fairly certain that some of the disturbance behaviors noted above are indicative that autonomic arousal is occurring as well (e.g., piloerection, retreat, cage clutch). In sum, we have clear indications that our monkeys are demonstrating fear in two of Lang's three response systems, with indirect evidence that fear is occurring in the physiological response system as well. This documented similarity between the symptomatology of human phobic fears and our monkeys' fear of snakes increases confidence that we have the potential for a valid and useful primate model of phobic fears. (For further discussion of the phobic quality of this fear of snakes, see the next section.)

THERAPY FOR FEAR OF SNAKES

PARALLELS IN TREATMENT OUTCOME FOLLOWING FLOODING THERAPY

As noted earlier there has been a long-standing controversy as to whether the fear of snakes observed in many primate species is innate or learned, with a preponderance of recent evidence strongly suggesting that it is learned. In contrast to the attention that has been paid to this issue, very little attention has been paid to the question of how easy it is to modify this fear. Schiller (1952) claimed that his chimpanzees' fear of snakes diminished quite rapidly when they were subjected to a counterconditioning-like procedure. However, his only index of the chimpanzees' fear was their latency to reach for a piece of food near the snake (i.e., he did not monitor any signs of behavioral distress or physiological arousal). Therefore it is not known whether other components of the fear were also reduced. Furthermore, he did not provide any long-term follow-up results and so there is no indication of the persistence of any extinction effects he found. Similar problems confound another published report of attempts to reduce the fear of snakes exhibited by wild-reared squirrel monkeys (Murray & King, 1973). These problems must be carefully considered given the research discussed earlier on desynchrony between measures of fear, especially during extinction (e.g., Grey, Sartory, & Rachman, 1979; Mineka, 1979), combined with those of Rachman (1979) on the return or spontaneous recovery of fear. Given such findings, neither Schiller's, nor Murray and King's (1973) results seem at all conclusive on the question of the persistence or resistance to extinction of snake fear in primates.

Some years ago we reported the results of an attempt to extinguish the intense fear of snakes exhibited by wild-reared rhesus monkeys (Mineka & Keir, 1983; Mineka *et al.*, 1980). All monkeys had met a criterion of showing an intense fear both by an index of behavioral avoidance and by an index of behavioral disturbance. They then received at least 12 sessions of a flooding-like procedure (seven in the first

month and five 6 months later). During each flooding session they were exposed to a live 3- to 4-foot boa constrictor (*Constrictor constrictor*) for a series of one to eight minute-long trials. Each trial ended when they had reached for their food treat on the far side of the snake (except that each trial had a minimum one minute trial duration). Sessions lasted until the monkey had reached a criterion of reaching for the food in less than 10 seconds on four consecutive trials. Consistent with the results of Schiller (1952) and of Murray and King (1973), we found that the monkeys all reached this criterion of four consecutive short-latency responses quite rapidly (none ever took more than 18 trials to do so).

Several aspects of our results were, however, strikingly different from the ones reported earlier. First, as seen in Figure 1, when the second flooding session occurred 10 days following the first session, there were no residual signs of reduced fear from the first session, that is, there was complete spontaneous recovery of fear. Sessions 2–5 occurred on consecutive days and a consistent trend toward improvement was observed on those sessions, that is, fewer trials to criterion across sessions and a shorter mean average latency per trial across sessions. Sessions 6 and 7 were each done at 10-day follow-up intervals, and during both sessions there was significant, although not complete, spontaneous recovery of fear. (See Figure 1). Six months later there was complete spontaneous recovery, with no residual signs of improvement. (See Mineka & Keir, 1983, for details.)

The second striking aspect of our results came from our observations of behavioral disturbance over the course of the flooding sessions. During the first seven sessions when the monkeys were showing a significant reduction in behavioral avoidance, there were no significant changes in behavioral disturbance. As seen in Figure 2 there were significant declines in behavioral disturbance from the first two to the last two trials of a session on several of the sessions, but there were no accompanying between-session changes. Furthermore, when further flooding sessions were conducted 6 months later (at the end of which there had been a total of 4 to 11 hours of exposure to the snake), there were still no significant changes in behavioral distress. Yet this is the component of the fear that we consider to be most akin to the subjective/verbal component in humans (see Mineka & Keir, 1983, for details).

Thus the subjective distress component of this fear of snakes exhibited by wild-reared monkeys appears to be highly persistent and resistant to substantial change. Indeed, in a very real sense one could reasonably argue that this fear is irrational in the sense used to describe phobic fears in the DSM-III (1980). The monkeys had safely reached for their food treat in the presence of the snake on many dozens of trials, and with very rapid latencies on at least the 48 criterion trials. In other words, they seemingly knew in the cognitive sense that this situation posed no real danger, and yet they continued to show undiminished levels of disturbance. This pattern of desynchrony or dissociation between different measures of fear closely parallels what has often been observed in human phobics undergoing flooding therapy. Indeed, Hodgson and Rachman (1974) noted that

it would appear that the first beneficial effect of flooding is an ability to control unwanted responses at the behavioral level. Autonomic and subjective signs of distress, associated with non-avoidance, are then gradually extinguished over a period of days, weeks, or months. (p. 321)

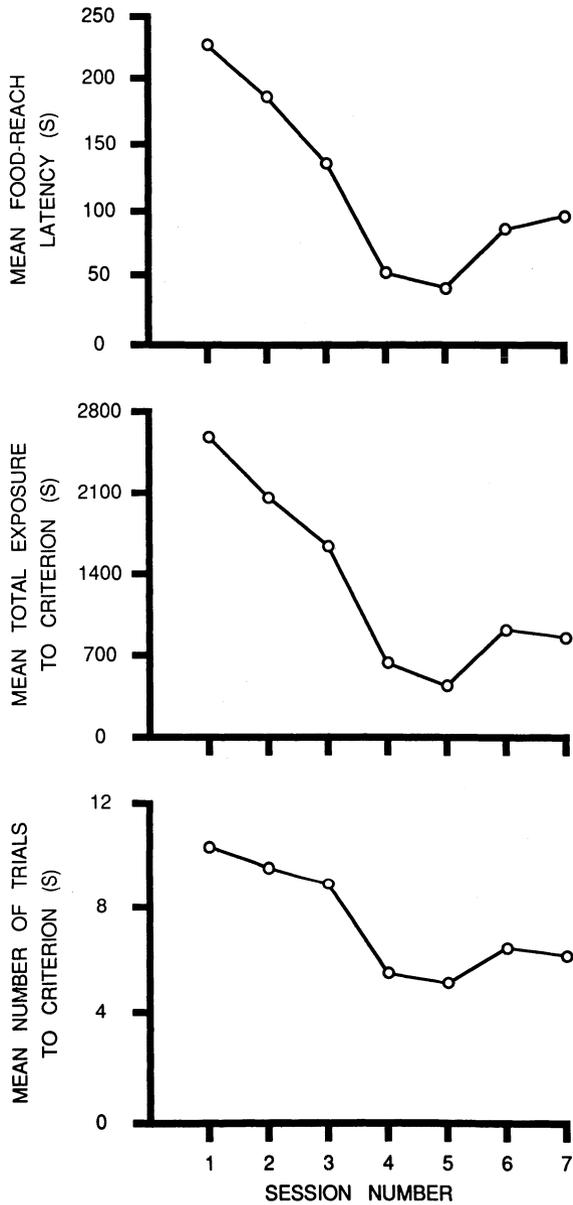


FIGURE 1. Mean latency, mean trials to criterion, and mean total exposure to criterion across the seven flooding sessions. There was a 10-day interval between Sessions 1 and 2, between Sessions 5 and 6, and between Sessions 6 and 7. (From Mineka *et al.*, 1980.)

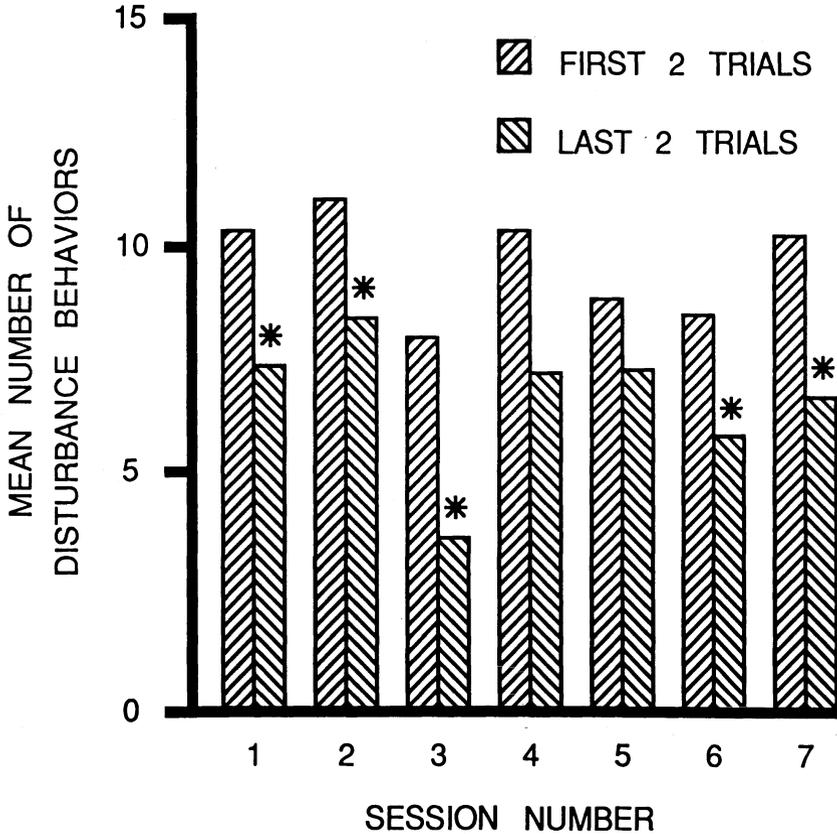


FIGURE 2. Mean behavioral disturbance scores for the first minute averaged across the first two and the last two trials of each of the seven flooding sessions. The asterisk (*) indicates that the first two vs. last two comparison is significant. (From Mineka *et al.*, 1980.)

Thus there are close parallels between the effects of flooding therapy on rhesus monkeys' fear of snakes and on humans' phobic fears. This similarity provides further support for the usefulness and validity of this primate model of phobic fears.

ETIOLOGY THROUGH OBSERVATIONAL CONDITIONING

Background. In recent years there has been increasing dissatisfaction with the classic theory originally proposed by Watson and Rayner (1920) that most human fears and phobias emerge through a process of direct traumatic classical conditioning. Sources of dissatisfaction with this theory are numerous, but primary among them are the observations that many people who report fears and phobias cannot recall any traumatic experiences having occurred in the presence of their now feared object. Observations of this type provide estimates ranging from 0% to 60% of intense fears or phobias having occurred as a result of direct traumatic conditioning (e.g., Emmelkamp, 1982; Murray & Foote, 1979; Öst & Hugdahl, 1981). Even in cases where

there is no recollection of a traumatic conditioning experience, social learning theorists have nevertheless been convinced that learning has played a role in the origins of phobic fears. In particular, they have speculated that observational or vicarious conditioning may account for the origins of a high proportion of human fears and phobias (e.g., Bandura, 1969; Marks, 1969; Rachman, 1977, 1978).

Until recently, however, the evidence to support such a proposal has been largely anecdotal or unconvincing. There are several dozen human studies documenting vicarious conditioning of autonomic responses such as heartrate and electrodermal responding (see Green & Osborne, 1985 for a recent review). However, the conclusions that can be drawn from these studies are quite limited for several reasons. First, none have employed anything other than autonomic indices of fear. Therefore, conclusions about whether fear, as indexed by the other fear-response systems, can also be vicariously conditioned are unwarranted. Second, because of ethical considerations that do not allow induction of severe and long-lasting fears in human subjects, none of these studies have tested for context specificity, persistence, or maintenance of the fear at long-term follow-up. Yet, by their very definition, the intense human fears and phobias whose origins are in question involve fear in multiple response systems. Furthermore, human phobic fears are not context specific, and they are both persistent and maintained over prolonged intervals in which the feared object is not encountered. (See Mineka *et al.*, 1984, for further discussion of these limitations.)

Basic Paradigm. More than a half dozen experiments from our laboratory have now convincingly documented rapid, strong, and persistent vicarious conditioning of snake fear. Such learning occurs when lab-reared monkeys that are not initially afraid of snakes are exposed to wild-reared monkeys behaving fearfully with snakes and nonfearfully with other objects. Like phobic fears in humans, this observationally learned fear is stimulus specific, but not context specific, and shows no signs of diminution at 3-month follow-up.

The procedure used to demonstrate this observational conditioning involves the Sackett Circus and the WGTA described earlier. All lab-reared observer monkeys are pretested in both situations to assure that they do not exhibit a fear of snakes. Model monkeys (usually wild-reared) are also pretested to assure that they do exhibit an intense fear of snakes. Following these pretests, the observer monkeys receive six discriminative observational conditioning sessions during which they watch the model monkeys in the WGTA behave fearfully with snakes and nonfearfully with neutral stimuli. Specifically, the observers are placed in a cage with a Plexiglas front several feet away from the model monkey. This allows them to observe the model monkey reaching or not reaching for food treats in the WGTA in the presence of a variety of stimulus objects placed in the open Plexiglas box. The observer monkeys can also clearly see the signs of behavioral distress exhibited by the model monkeys when they are reacting to the snake stimuli.

Each session consists of fifteen 40-second trials, six of which are with snake stimuli (real and toy) and nine of which are with neutral stimuli. Following two sessions of observational conditioning in the WGTA (involving a total of 8 minutes of exposure to the models behaving fearfully with snake stimuli), the observers are tested for fear of snakes by themselves in a different context—the Sackett Circus. After this first posttest in the Circus, the observer monkeys receive four more sessions

of observational conditioning in the WGTA, with additional Circus posttests occurring following the fourth and sixth observational sessions. Following the final Circus posttest, all observers receive a test by themselves in the WGTA to assess both behavioral distress and behavioral avoidance components of their acquired fear. Approximately 3 months later, all observers are given a follow-up test in both the Circus and the WGTA to assess retention of their acquired fear.

In our first study (Mineka *et al.*, 1984), the observers were six lab-reared adolescent/young adult rhesus monkeys (3 to 6 years of age) who had been reared with their wild-reared parents; the parents (father or mother) served as models. Five out of the six adolescent/young adult observers acquired an intense and long-lasting fear of snakes. This fear was exhibited at asymptotic intensity during their initial Circus test after only 8 minutes of watching one of their parents behaving fearfully with snakes. The fear manifested itself in both the Circus and the WGTA, through measures that tapped both the behavioral avoidance and the behavioral distress component of the fear. During the 3-month follow-up test, there were no signs that the acquired fear of snakes had diminished in intensity.

In the second study of the series, we explored whether the parent-child relationship that was an inherent part of the model-observer relationship in the first study was necessary to produce such rapid, strong, and persistent observational conditioning. The two models and the 10 observers in this study (Cook, Mineka, Wolkenstein, & Laitsch, 1985) were completely unrelated monkeys that were merely "acquainted" with one another by virtue of having lived in the same room together (not the same cage) in the recent past. The results produced with unrelated models and observers were highly similar to those produced in the first study when the observers were the offspring of the models. As can be seen in Figure 3, in the Circus pretest the models spent nearly all of their time in the neutral stimulus compartment, and virtually no time in the snake compartments. The observers, by contrast, spent comparable amounts of time in the neutral and snake compartments during the pretests. During the Circus posttests, however, the observers' choice of stimulus compartments closely paralleled that of the models, with very little time being spent in the snake compartments and a good deal of time being spent in the neutral compartment. (Substantial learning occurred in 7 out of 10 observers in this study, as compared with 5 out of 6 in the first study).

Similarly, as can be seen in Figure 4, during the WGTA pretests the models exhibited their fear by showing very long (usually maximal) latencies to reach for food only in the presence of snake stimuli. During pretesting the observers exhibited a lack of snake fear by showing short latencies in the presence of both snake and neutral stimuli. In the WGTA posttest, the observers' pattern of behavior closely paralleled that of the models, with long food-reach latencies now being shown to the snake stimuli, but with short latencies to the neutral stimuli still being maintained. Finally, as seen in Figure 5, the results for the behavioral distress component of the fear closely parallel those for the behavioral avoidance component. Models showed high levels of behavioral disturbance only in the presence of snake stimuli. During the pretest observers showed low levels of disturbance in the presence of all stimuli, but during the posttest they showed high levels of disturbance in the presence of

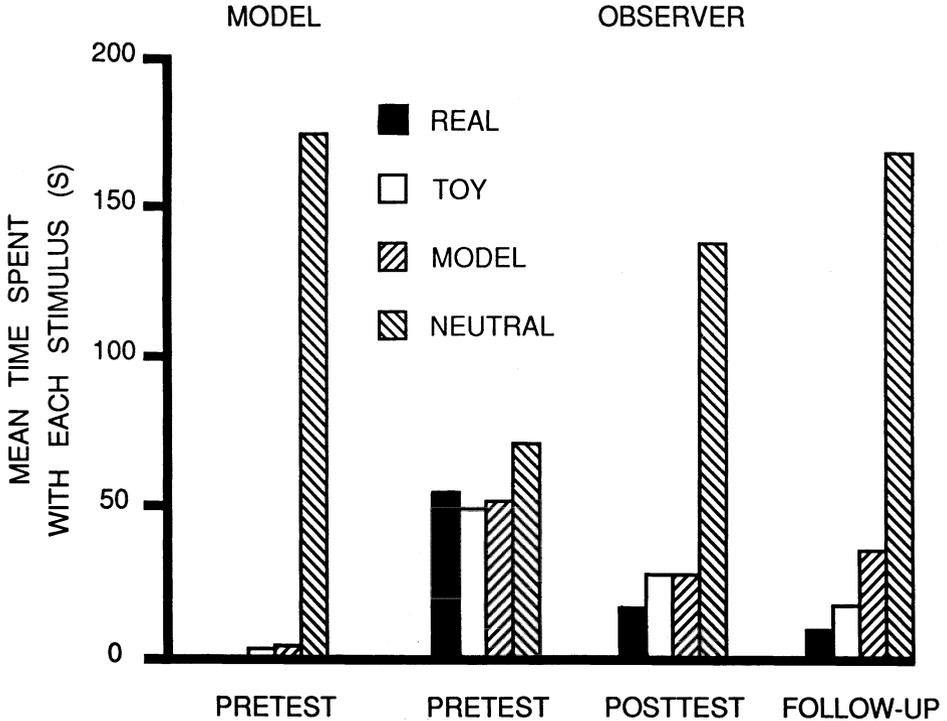


FIGURE 3. Mean amount of time spent with the four different objects in the Sackett Circus for the models in the pretest, and for the observers in the pretest, posttests, and three month follow-up. (From Cook *et al.*, 1985.)

snake stimuli. It should also be noted in each of these figures that the results indicated no signs of a decrease in intensity of the acquired fear at 3-month follow-up.

Individual Differences. Four of the 16 subjects in these first two experiments did not show very significant conditioning. This raises the interesting question of what is the source of the differences between the four who did not learn and the 12 who did learn? One possibility is that the four who did not learn were simply bolder and less emotional monkeys who would show weaker conditioning in any fear conditioning paradigm. Alternatively, the observational conditioning experiences of the four subjects who did not learn may somehow have been different from those of the 12 subjects who did learn. The second possibility must be seriously considered because the experimenter cannot control the fearful model's behavior during the observational conditioning sessions. Therefore the observational conditioning experiences of the observers vary across sessions and across observers in ways they do not in traditional classical conditioning paradigms where the experimenter controls the presentation of the unconditioned aversive stimulus.

One can begin to explore the source of the individual differences in conditioning by examining the relationship between the models' level of fear exhibited during conditioning and the observers' level of fear in the posttests. The models' behavior

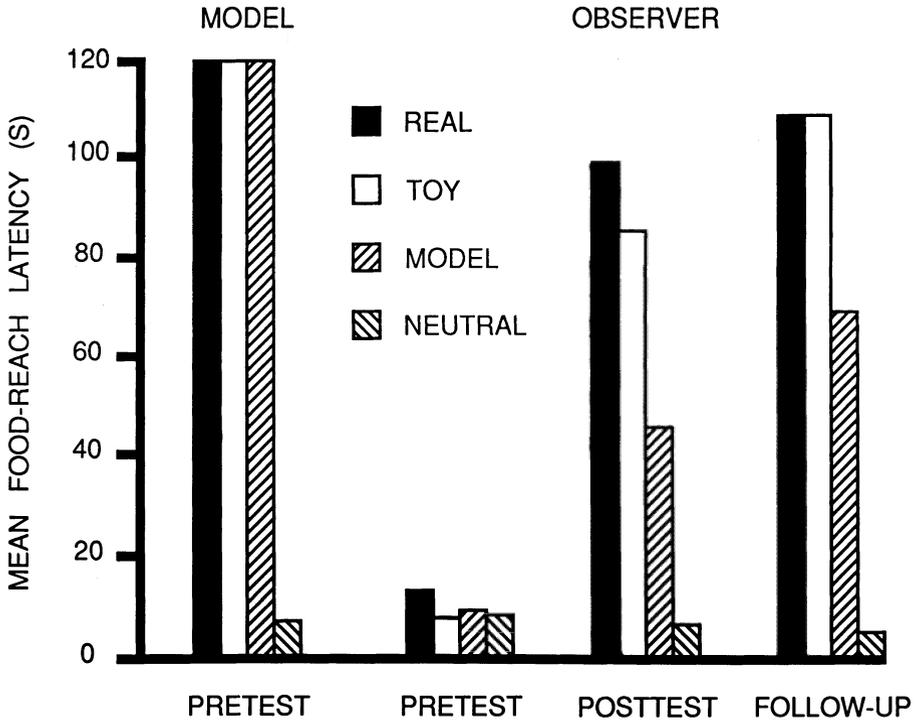


FIGURE 4. Mean food-reach latency in the Wisconsin General Test Apparatus in the presence of the four different objects for the models in the pretest, and for the observers in the pretest, posttest, and three month follow-up. (From Cook *et al.*, 1985.)

during the observational conditioning sessions (food-reach latency and levels of behavioral disturbance) had been recorded, and so it was possible to correlate the level of fear that they exhibited with that of the observers in the posttest. In both experiments described above (Mineka *et al.*, 1984, Experiment 2; Cook *et al.*, 1985, Experiment 1), there were very high correlations (r 's = .986 and .95, respectively) between the total amount of disturbance behaviors exhibited by the models to the three snake stimuli *during* conditioning, and by the observers in the WGTA posttest. Such high correlations in both experiments suggests that observers may closely model the degree of fear or disturbance that a particular stimulus elicits in the model. An alternative explanation of these high correlations stems from a possibility, discussed in the next section, that the model's fear display serves as a kind of unconditioned stimulus eliciting an unconditioned response of distress in the observer. By this view these high correlations may simply reflect the fact that superior conditioning generally occurs with stronger unconditioned stimuli. In either case, a large part of the answer to the question of the source of the marked individual differences in observational learning seems to lie in the differences in the observational conditioning experiences that the different observers receive. One interesting implication of these findings concerns possible ways of reducing the impact of vicarious conditioning experiences.

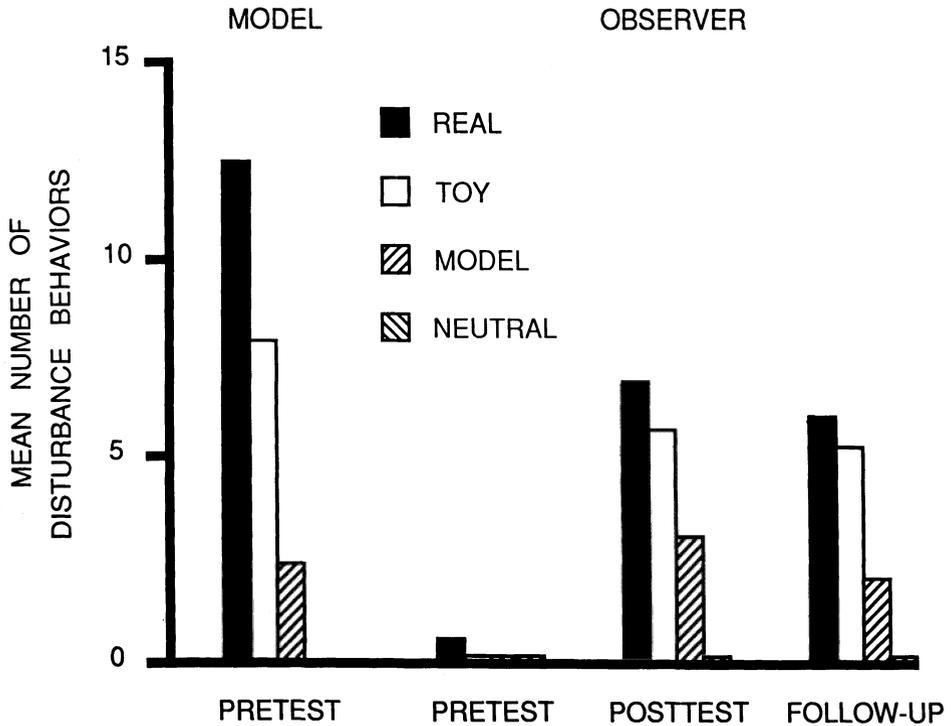


FIGURE 5. Mean number of disturbance behaviors exhibited in the Wisconsin General Test Apparatus in the presence of the four different objects for the models in the pretest, and for the observers in the pretest, posttest, and three month follow-up. (From Cook *et al.*, 1985.)

In particular, it suggests that parents who have strong fears or phobias should attempt, while in the presence of their children, to minimize their level of distress if they encounter their phobic object. More intense fear displays should be associated with a greater likelihood of the children (or others) vicariously acquiring the modeled fear.

Possible Mechanisms Underlying Observational Conditioning. As discussed earlier, there was a strong relationship in our first two studies between a model's level of fear exhibited during conditioning and the observer's level of acquired fear during the posttest. In addition, the observers also showed significant levels of disturbance during conditioning as they watched the model's fear display. Indeed, their own levels of fear or disturbance during conditioning were also highly correlated with the levels of fear exhibited by the model, and with their own level of acquired fear during the posttests.

This raises the interesting question of the nature of the observer's distress reaction to seeing the model's fear display. Is it like an unconditioned response to an unconditioned stimulus (the model's fear display)? Or does something more akin to a cognitive social inference process occur in which the observer sees the model's response to the snake stimuli, and infers that she or he too should be afraid? Although

these two possibilities may be difficult to tease apart in a definitive way, they do indeed seem to be distinct. As a first attempt to tease them apart, we reasoned as follows: if the observer's response is like an unconditioned response (UCR) to the model's fear display (UCS), then the observers should show reactions of comparable intensity whether or not they can see what the model is reacting to. Alternatively, if a process more akin to cognitive social inference occurs, then observers not able to see what the model is reacting to should show less distress than observers who can see the snake stimuli to which the model is reacting.

In order to examine this question experimentally, we have recently completed a study in which half of the observers were only able to see the model's fear display, and not what the model was reacting to (Mineka & Cook, 1987). The other half of the observers were able to see the snake stimuli as well as the model's reactions to them. During six such sessions that were otherwise highly similar to those described earlier for the basic paradigm, the observers' reactions to the model's fear display were carefully observed. The results, illustrated in Figure 6, were quite intriguing. During the first session, observers in both groups showed comparable levels of fear while watching the model's fear display. However, during subsequent sessions (2-6) only the observers that could see the snake stimuli to which the models were reacting continued to show signs of fear; observers that could only see the model but

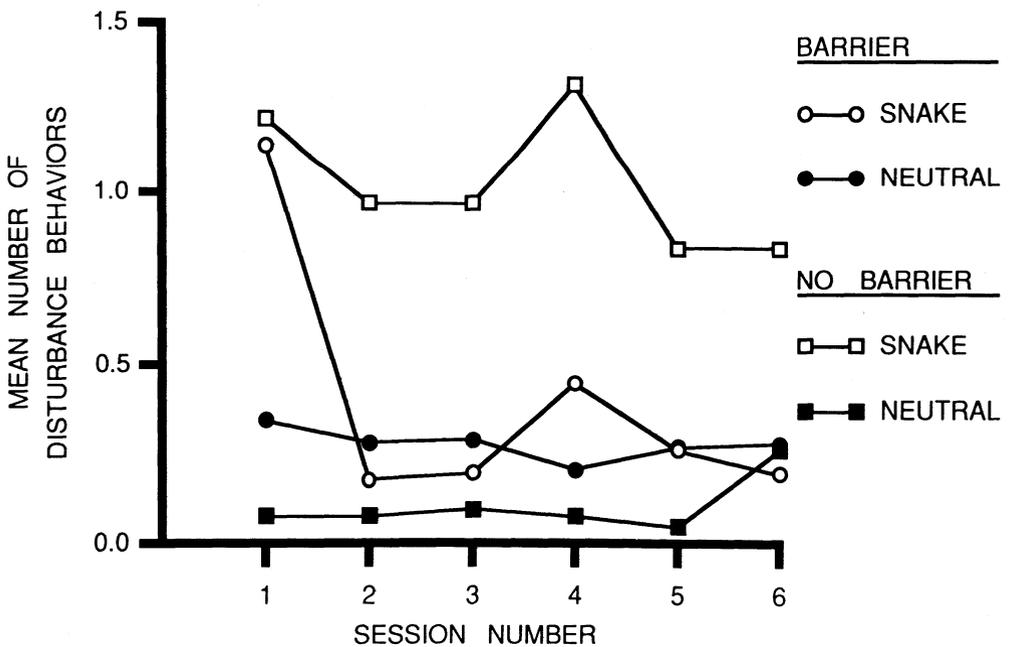


FIGURE 6. Mean number of disturbance behaviors exhibited by the observers across sessions in response to the model's fear display to snake stimuli. The no barrier group could see the model and the snake stimuli eliciting fear in the model; the barrier group could only see the model's fear display and not the snake stimuli. (From Mineka & Cook, 1987.)

not what she or he was reacting to no longer showed visible signs of disturbance. This suggests that the observer's fear response to the model's fear display may be, at the outset, akin to a UCR to a UCS. However, it appears to be a somewhat fragile UCR that habituates rather rapidly, as evidenced by the virtually nonexistent levels of disturbance in the observers not able to see the snake stimuli on Sessions 2 through 6.

Why then do the observers who can see the snake stimuli continue to manifest comparable levels of fear across all six sessions? The most plausible answer to this question seems to be that by the second session these observers have already acquired a fear of snakes. In other words, their continued high levels of fear seem to be an indication of their own level of acquired fear, rather than a reaction to the model's fear display *per se*. This account seems plausible because we have independently demonstrated in another study that significant levels of snake fear are acquired after only one session of discriminative conditioning (Mineka & Cook, 1987).

Another intriguing question concerning the observers' reactions to the models' fear displays is whether they are necessary for conditioning to occur. As discussed earlier, at the outset the observers' reactions appear to be akin to a UCR to a UCS (the model's fear display). However, what is still unclear is whether the observer's disturbance to the model's fear response plays a critical role in mediating the learning. That is, is the observer's disturbance necessary to the learning process, or is it a mere concomitant of the learning process? One way to begin to answer this question would be to examine the effects of administering various antianxiety drugs to the observers prior to their observational conditioning sessions. It is possible that under the effects of a benzodiazepine or a beta-blocker, an observer's fear reaction to a model's fear display might be dampened and yet learning might still occur. Such results would suggest that the observer's fear reaction was not a necessary part of the learning process, but rather a mere concomitant of it. Such a study has not yet been completed, but seems to be an important one for future research.

The Belongingness Issue and the Nonrandom Distribution of Fears and Phobias. In recent years numerous theorists of fears and phobias have noted that the objects of most people's fears and phobias do not generally come from a random arbitrary group of objects (e.g., Marks, 1969; Rachman, 1978; Seligman, 1971). For example, people do not usually develop strong fears of electric outlets, hammers, bicycles, stoves, etc., even though such objects may frequently be associated with trauma and/or with verbal instructions to avoid these objects. Instead, most people tend to develop strong fears or phobias about snakes, spiders, water, heights, etc. Thus, it seems likely that fears of some objects may be more easily acquired and/or more difficult to extinguish than are fears of other objects.

In the past decade Öhman and his colleagues (e.g., Öhman, 1986; Öhman, Fredrikson, Hugdahl, & Rimino, 1976; Öhman, Fredrikson, & Hugdahl, 1978; Öhman, Dimberg, & Öst, 1985) have reported the results of an elegant series of experiments demonstrating a number of different characteristics of conditioning to "fear-relevant" as opposed to "fear-irrelevant" stimuli. Using mild electric shock as a US and electrodermal responses as an index of fear in nonphobic human subjects, Öhman *et al.* have found that conditioned responses (CRs) to common phobic stimuli (e.g., slides

of snakes, spiders, and angry faces) extinguish only very slowly. Consistent with the irrationality of phobic fears, such CRs are also not affected by cognitive instructional variables, such as informing the subject that shocks will no longer occur. CRs established to fear-irrelevant stimuli (e.g., slides of geometric objects, flowers, or happy faces), on the other hand, extinguish quite rapidly and are sensitive to cognitive instructional variables.

In addition, Cook (1983; see also Cook, Hodes, & Lang, 1986) demonstrated that it is essential to use a tactile US such as electric shock to obtain such effects; a loud noise US does not produce parallel results. Such results are of importance in the demonstration of selective associability, that is, snakes and spiders are not simply superior CSs in general, but rather only when they are associated with certain types of USs. Cook also found evidence that the heartrate CR to fear-relevant CSs was acceleratory in nature rather than deceleratory as with fear-irrelevant CSs. Thus, there may be more of a defensive component in the CRs to fear-relevant CSs, and more of an orienting component in the CRs to fear-irrelevant CSs. This general line of experimentation has been extremely important in revitalizing interest in conditioning models of phobias. As pointed out by a number of theorists, such models had never been especially compelling precisely because of the nonarbitrary nature of the stimuli involved in phobias, their very high resistance to extinction, the irrationality of phobic fears, etc. (e.g., Mineka, 1985; Rachman, 1977, 1978; Öhman *et al.*, 1978; Öhman *et al.*, 1985; Seligman, 1971).

One prominent theory that has been proposed to account for the superior conditioning seen with fear-relevant stimuli is Seligman's (1971) preparedness theory. According to this theory, we are evolutionarily "prepared" to associate with aversive or traumatic events certain stimuli that were often dangerous or threatening to our early ancestors. In this view, those ancestors who acquired fears of these objects easily, and who maintained them for prolonged periods of time, may have had a selective advantage in the struggle for existence. This selective advantage would have been in comparison to their contemporaries who did not acquire these fears so easily, or for whom the fears, once acquired, extinguished quite rapidly.

Öhman *et al.* (1985) recently proposed a more complex evolutionary model distinguishing between the evolutionary pressures that may have led to the easy acquisition of animal or interspecific fears (such as snakes and spiders) versus social or intraspecific fears. They argue that animal or interspecific fears are highly reflexive and automatic in their elicitation and are "tightly organized as an escape or avoidance package." Social or intraspecific fears, by contrast, are "much more loosely and conditionally concocted, with a less prominent and reflexive role for active avoidance behavior" (p. 141). They amass a wide range of evidence from the human and animal conditioning literature, as well as from studies of phobic patients, to support their evolutionary hypotheses, which are considerably more specific and predictive than Seligman's original statement of the preparedness theory of phobias.

One of the issues that has created controversy for the preparedness theory of phobias concerns whether the superior conditioning seen with fear-relevant stimuli stems from evolutionary as opposed to ontogenetic factors (cf. Delprato, 1980; Mackintosh, 1974; Schwartz, 1974). In other words, the reason people show superior fear

conditioning with snakes or spiders may not reside in their evolutionary past, but rather in the associations they have acquired to these stimuli in their lifetime. Such associations are probably most often at least somewhat negative or unpleasant (although subjects do not have measurable levels of fear prior to their participation in an experiment). There is some research supporting the importance of biological as opposed to experiential factors in producing the pattern of differences described above. For example, Hugdahl and Kärker (1981) found superior conditioning with snakes and spiders, but not with electric outlets. This is in spite of the fact that electric outlets are presumably as likely to have negative associations to them built up during ontogeny as are snakes and spiders. Others have criticized these results, however, by noting that human subjects are also likely to have had a good deal more neutral (nonaversive) experience with electric outlets than with snakes and spiders. This difference in prior nontraumatic experience may account for differences in conditionability.

One thing seems certain from all this controversy about biological versus experiential contributions to the so-called preparedness effects: obtaining a definitive answer to this issue exclusively through research with human subjects is extremely unlikely because of the impossibility of controlling for their prior exposure to stimuli before participation in an experiment. An obvious advantage, then, can be seen to studying this question in laboratory-reared monkeys who, prior to our work with them, have had no previous exposure to snakes. By having such control over ontogenetic factors, one can make more informed inferences about the possible contribution of evolutionary factors to these effects.

In a preliminary experiment designed to begin to explore this question, we chose an overshadowing design in which naive observer monkeys were exposed to model monkeys reacting fearfully to a compound snake/flower stimulus (Cook & Mineka, 1987a). The brightly colored artificial flowers were placed on alternate trials in front of, or behind the snake stimuli. (To the best of our knowledge there was no way in which the observers could ascertain which part of the compound stimulus the model monkeys were reacting to.) Half of the monkeys were exposed only to toy snakes (no real snake) in order to control for possible differences in salience between the flower and real snake stimuli due to animateness or movement cues. The discriminative observational conditioning procedure for this experiment was identical in every other way to that in the previous experiments described earlier. The posttest procedure was modified only slightly by the addition of the flower stimulus to the posttests in the Circus and the WGTA.

The results of this experiment, illustrated in Figures 7 and 8, clearly show that the observers acquired a fear of the snake stimuli but not of the flower stimuli. Interestingly, there was some slight initial reaction to the flower stimuli in the model monkeys, probably a reaction to the novelty of this brightly colored display. Such a novelty reaction is also sometimes seen by naive observer monkeys when they first see snake stimuli. However, in both cases the slight hesitancy seen with the novel stimulus generally habituates quite rapidly, unlike the pronounced fear reactions that are acquired during observational conditioning. It is also important to note that the failure of acquisition of flower fear occurred both in monkeys for whom the flowers

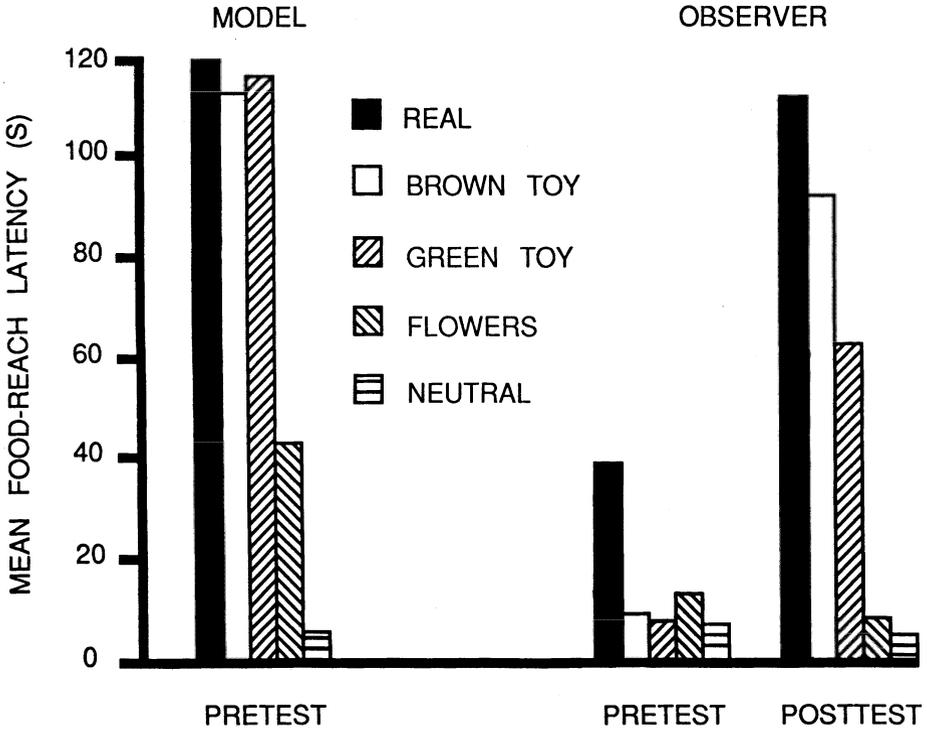


FIGURE 7. Mean food-reach latency in the Wisconsin General Test Apparatus in the presence of the five different objects (snakes, flowers, and neutral block) for the models in the pretest, and for the observers in the pretest and posttest. (From Cook & Mineka, 1987a.)

were compounded with a real snake stimulus and in monkeys for whom the flowers were compounded only with toy snake stimuli. Thus differences in salience or animateness do not appear to be able to account for these results. This seems especially likely given that the brightly colored artificial flowers are considerably more salient, at least to the human eye, than are the rather dull colored toy snakes.

Nevertheless, there are some limitations in the design of this experiment that do not allow firm conclusions to be drawn about the preparedness or belongingness issue discussed earlier. For example, without the inclusion of a group of observers that saw models reacting fearfully only to flowers, we cannot be certain about the difficulty of associating flowers with fear *per se*. Why was such a group not included in the experiment? Reflecting back to the earlier discussion on the sources of individual differences in acquired fear, the difficulty becomes clear. The major determinant of how much fear an observer acquires is how much fear his or her model exhibited during the conditioning sessions, with these correlations being at least .95 in our first two experiments. Thus, in order to demonstrate belongingness or preparedness (that is, superior acquisition of snake fear as compared to flower fear), one must be able to precisely equate the model's fear performance in the presence of flowers with his or her fear performance in the presence of snakes. Yet it would seem very difficult,

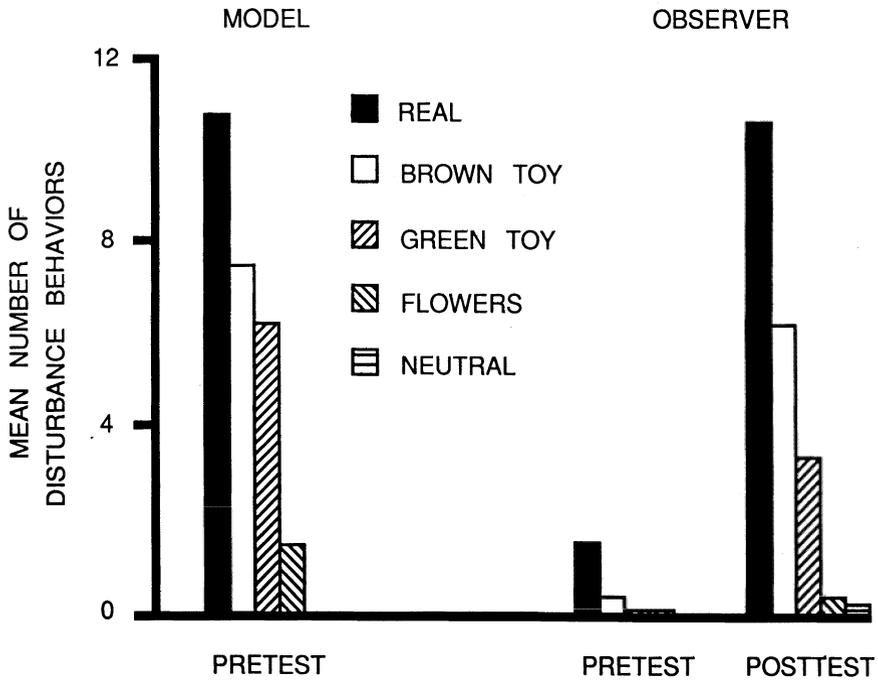


FIGURE 8. Mean number of disturbance behaviors exhibited in the Wisconsin General Test Apparatus in the presence of the five different objects (snakes, flowers, and neutral block) for the models in the pretest, and for the observers in the pretest and posttest. (From Cook & Mineka, 1987a.)

if not impossible to condition a fear of flowers in our model monkeys such that they behaved with exactly the same intensity of fear that they show with snakes.

The wonders of modern video technology have finally given us the needed tool to begin to explore this question in a more definitive way. Previous research by Capitanio, Boccia, and Colaianna (1985) had shown that young monkeys respond with socially appropriate reactions to videotapes of fearful/submissive monkeys, as well as to videotapes of monkeys exhibiting a threat. This led to the hypothesis that perhaps monkeys could also profit from the experience of watching videotapes by, for example, learning fears that have been modeled by other monkeys on the videotape. If such learning could be demonstrated through the use of videotapes, then we would be able to begin to explore the belongingness issue by editing the videotapes. Specifically, editing/splicing techniques would allow us to equate a model's fear display with various stimuli that vary in fear relevance, for example, snakes and flowers.

Initially it was necessary to investigate whether naive observer monkeys would indeed acquire a fear of snakes simply through watching a color videotape of a model monkey reacting fearfully to snakes and nonfearfully to non-snake stimuli. Videotapes were made of two different wild-reared monkeys during the equivalent of two discriminative observational conditioning sessions like those described above. (See Cook & Mineka, 1987b, for details). These videotapes were subsequently shown to six

naive observer monkeys. In order to maximize the possibility of obtaining conditioning, each observer watched two videotaped sessions a day (one of each model) for six days. This resulted in twice as much exposure to fearful models as in our traditional live observational conditioning procedure.

As can be seen in Figures 9 and 10, the results were strikingly similar to those obtained with the use of live models (compare with the results illustrated in Figures 3–5). In particular, rapid, strong, and persistent observational conditioning of snake fear was exhibited during the posttests. In addition to being of importance for pursuing the belongingness issue, these results are of interest in their own right. Bandura and his colleagues documented many years ago that human observers who watch aggressive behavior being modeled, live or on videotape, have a marked tendency to later display the same or related aggressive behaviors. (See Bandura, 1969, for a review). Many have also speculated that humans may acquire fears of previously unfeared objects simply through watching television, but the evidence documenting such proposals has been primarily anecdotal. The previously cited results, however, provide a clear-cut demonstration that monkeys, and probably humans as well, can indeed acquire a long-lasting and intense phobic-like fear simply through watching fearful models on videotape (television).

In order to pursue the belongingness issue, new videotapes were made that had trials of three different types: (a) neutral stimulus (wood block) trials on which the model reacts nonfearfully to a wood block; (b) snake trials on which the model reacts

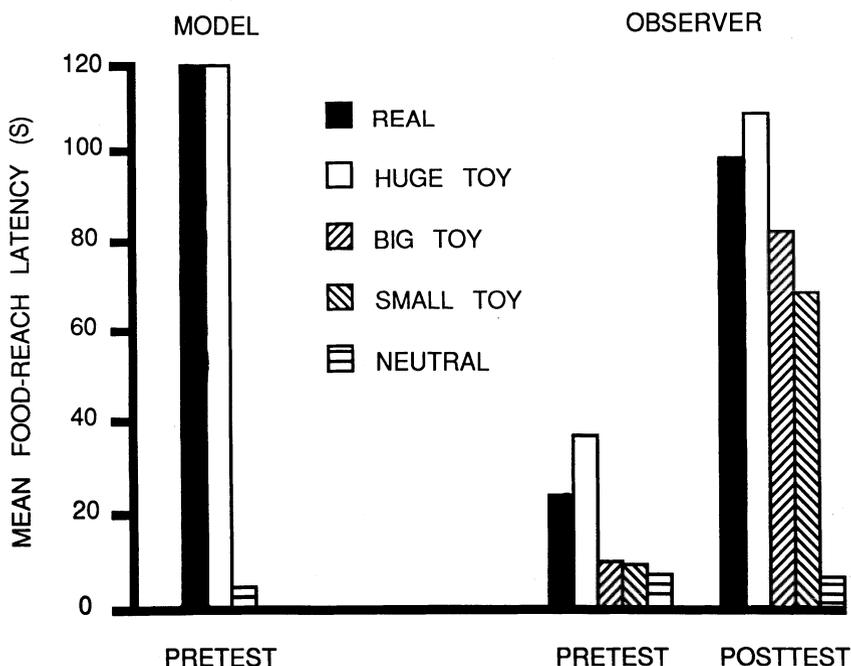


FIGURE 9. Mean food-reach latency in the Wisconsin General Test Apparatus in the presence of the five different objects for the models on the videotape, and for the observers during the pretest and posttest. (From Cook & Mineka, 1987.)

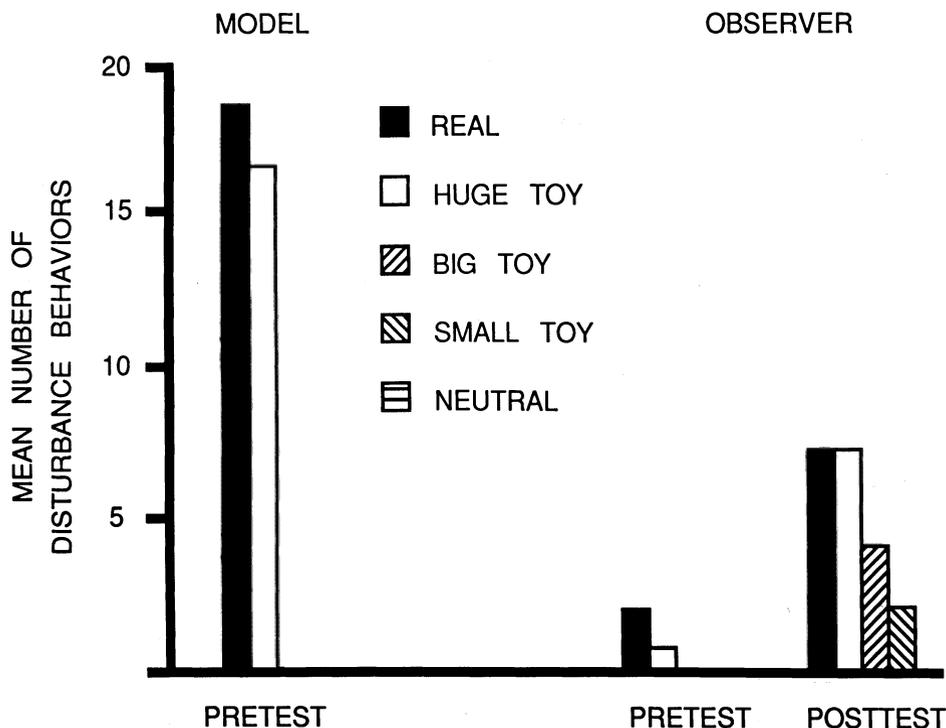


FIGURE 10. Mean number of disturbance behaviors exhibited in the Wisconsin General Test Apparatus in the presence of the five different objects for the models on the videotape, and for the observers during the pretest and posttest. (From Cook & Mineka, 1987.)

fearfully to a live snake; and (c) flower trials on which the model reacts nonfearfully to brightly colored artificial flowers. Then two different versions of *edited* videotapes were made. On the SN+/FL- version, CS+ (excitatory) trials consisted of a monkey, who in reality had been reacting fearfully to a live snake, appearing to react equally fearfully to the spliced-in image of a toy snake. CS- (inhibitory) trials for this SN+/FL- videotape consisted of the monkey reacting nonfearfully to the spliced-in image of different brightly colored flowers. On the second FL+/SN- version of the edited videotape, CS+ trials consisted of a monkey, who in reality had been reacting fearfully to a live snake, appearing to react fearfully to the spliced-in image of brightly colored artificial flowers. CS- trials for this second FL+/SN- videotape consisted of the monkey reacting nonfearfully to the spliced-in image of a toy snake. Thus, for the CS+ trials the two different versions of the videotape equate the model's fear performance with toy snakes and with brightly colored artificial flowers. Indeed, the exact same fear display footage is used for the CS+ trials on each videotape, and the exact same non-fear-display footage is used for the CS- trials on each videotape. Consequently the total amount of fear displayed to the toy snakes on the SN+/FL- tape is identical to the total amount of fear displayed to the brightly colored flowers on the FL+/SN- videotape. Furthermore, because the edited versions of the videotape use only toy snake stimuli, in relatively drab colors, there are

no apparent differences in animateness or salience between the fear-relevant and fear-irrelevant cues. Indeed, if anything the brightly colored artificial flowers are far more salient to the human eye on the color videotape than are the toy snake stimuli. Thus it does not seem that differences in salience or animateness could account for any superior conditioning to the fear-relevant stimuli that might occur. (See Cook & Mineka, 1987b, for details).

Preliminary results of this experiment, which is still in progress, support the preparedness/belongingness hypothesis. Monkeys exposed to the SN + /FL - videotapes have shown significant acquisition of snake fear, with no sign of acquiring a fear of flowers. (Overall levels of acquired snake fear were somewhat lower than in the previously described videotape experiment, presumably because of the use of less salient toy snake stimuli rather than a live snake.) Figures 11 and 12 illustrate this discriminative conditioning of snake fear in the SN + /FL - group. By contrast, monkeys exposed to the FL + /SN - videotapes have for the most part failed to acquire a fear of either snakes or flowers. This failure to acquire a fear of flowers is in spite of the fact that *exactly the same fear displays* were exhibited to the brightly colored flower

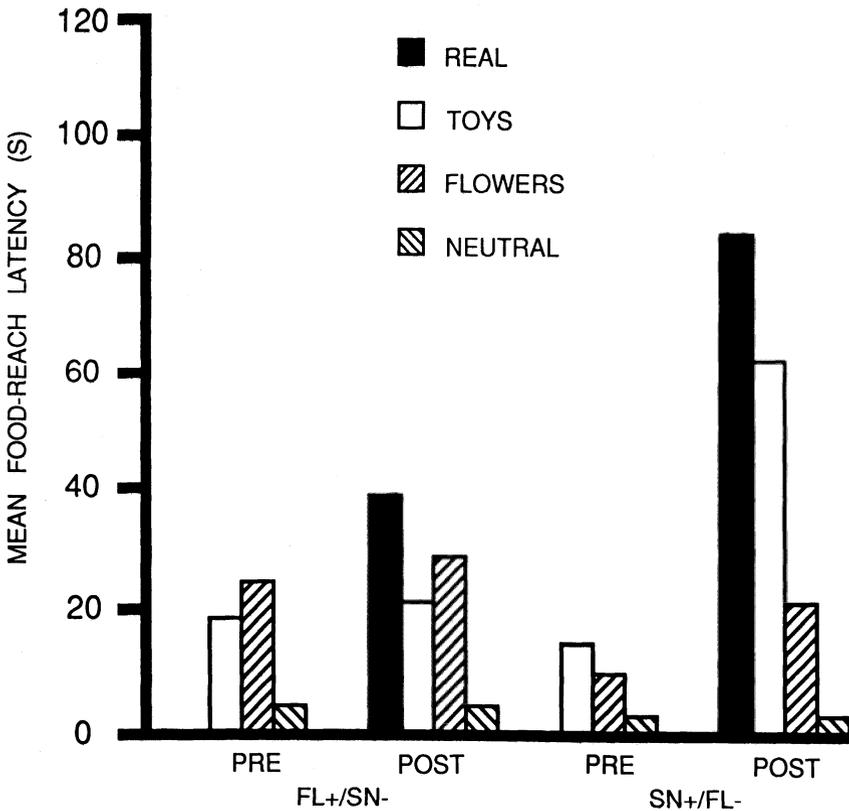


FIGURE 11. Mean food-reach latency in the Wisconsin General Test Apparatus in the presence of the four different objects during the pretest and posttest for the observers in the SN + /FL - group and the FL + /SN - group. (From Cook & Mineka, 1987.)

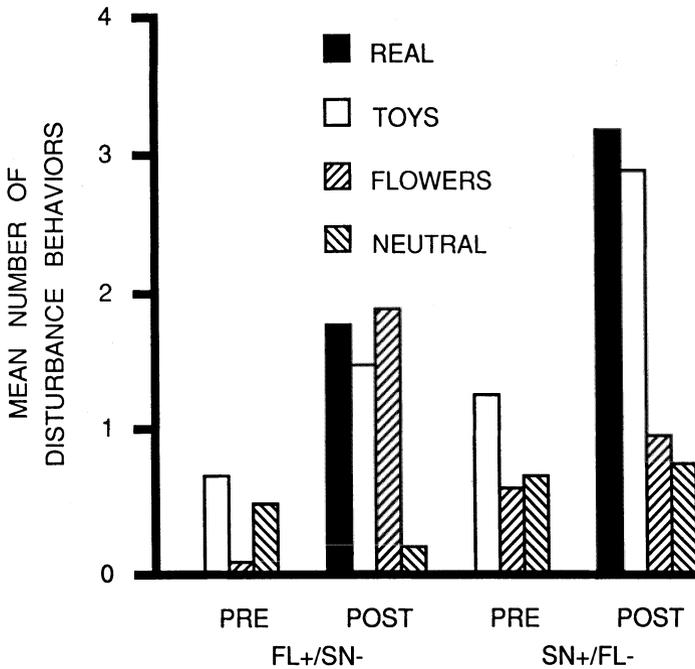


FIGURE 12. Mean number of disturbance behaviors exhibited in the Wisconsin General Test Apparatus in the presence of the four different objects during the pretest and posttest for the observers in the SN+/FL- group and the FL+/SN- group. (From Cook & Mineka, 1987.)

stimuli on the FL+/SN- videotape as were exhibited to the toy snake stimuli on the SN+/FL- videotapes.

Such results provide quite strong evidence for an evolutionarily based belongingness between certain kinds of common phobic stimuli and fear. This conclusion will be even more certain if it can also be demonstrated that these brightly colored flower stimuli that failed as CS+s for fear are not simply inadequate stimuli for the conditioning of any type of response. That is, it will be important to show that they can be used to condition some type of appetitive response before we conclude with certainty that snake stimuli are differentially associable with fear as opposed to simply being more salient and conditionable CSs in general. Future research is planned to explore this issue.

Summary and Implications of Observational Conditioning of Fear for the Origins of Human Fears and Phobias. The experiments described earlier have convincingly demonstrated that intense and phobic-like fears can be learned through observation alone after relatively brief exposure to a fearful model. The model's level of fear during conditioning strongly determines the level of fear that the observer acquires. These convincing empirical demonstrations using a primate model provide strong support for the hypotheses of clinical researchers who increasingly have attributed an important role to observational conditioning in the origins of human phobic fears. Previously such hypotheses had only been supported by laboratory demonstrations in humans of vicariously acquired heartrate and electrodermal responses and by retrospective

reports of the onset of human phobias. That the fear can be acquired equally easily through observation of related and unrelated models, live or on videotape, also documents the potentially wide variety of situations in which such fears can be learned. Finally, that the observationally learned fear is acquired more easily to some fear-relevant stimuli than to fear-irrelevant stimuli, provides a good model for understanding the nonrandom distribution of human fears and phobias.

IMMUNIZATION/PREVENTION OF FEARS

If humans show as rapid acquisition of phobic fears through observation as do monkeys, it is perhaps somewhat surprising that the incidence of specific fears or phobias is not especially high in friends and relatives of individuals exhibiting intense specific fears or phobias. Given the numerous opportunities children have to see their parents exhibit their fears, one might be especially likely to expect a higher concordance rate between parent's and children's fears than has typically been reported. (See Emmelkamp, 1982; Marks, 1987, for reviews.) There is, however, one important major difference between the observer monkeys used in the experiments described and many humans who observe models behave fearfully with specific objects. In particular, the observers used in these experiments had had very little exposure to snake stimuli prior to their observational conditioning experiences (approximately 13 minutes during the pretests). By contrast, many humans will have had much more extensive prior exposure to an object before they see a model behaving fearfully with that object. Such prior experience with an object can occur in two ways. First, the person may spend time alone with the object, thereby increasing the familiarity of the object and perhaps reducing the object's salience (a phenomenon called latent inhibition in the conditioning literature, cf. Mackintosh, 1974, 1983). Alternatively or additionally, people may have been exposed to another person behaving nonfearfully with the object (i.e., a nonfearful model). Either or both of these types of prior exposure might be expected to serve as sources of immunization against the effects of subsequent observational conditioning experiences. Unfortunately, however, there are no empirical demonstrations in humans that such immunization effects can indeed occur.

Again, using a primate model, we have recently demonstrated that one of these types of prior exposure—observation of a nonfearful model behaving nonfearfully with snake stimuli—can successfully prevent subsequent observational conditioning (Mineka & Cook, 1986). The immunization procedure used in this experiment lasted six sessions. It was identical to the discriminative observational conditioning procedure described earlier using snake and non-snake stimuli, except that the models were laboratory-reared monkeys who were not afraid of snakes. These immunization sessions were followed by the traditional six sessions of discriminative observational conditioning with a fearful model behaving fearfully with snakes. Two other comparison groups that differed in what happened in the first phase of preexposure were also included: (a) a pseudoimmunization control group that first spent six pseudoimmunization sessions watching nonfearful models behaving nonfearfully with neutral stimuli, and (b) a latent inhibition control group that spent six sessions by themselves

with snake and non-snake stimuli (no models were present). The latent inhibition group was equated with the immunization group for the total amount of preexposure to snake stimuli that they had had. These two groups differed in whether that exposure to snake stimuli had occurred when they were by themselves, or when they were watching a nonfearful model behaving nonfearfully.

The results revealed large and significant group differences when all three groups were tested for acquisition of snake fear following their six observational conditioning sessions with fearful models. As illustrated in Figures 13 and 14, the pseudoimmunization group showed high levels of acquired fear comparable to that seen in our prior experiments. By contrast, the immunization group did not show significant changes in their responding to snakes from pretest to posttest, and showed significantly less fear by indices of behavioral disturbance and behavioral avoidance than did the pseudoimmunization group. The results of the latent inhibition group were intermediate and not generally significantly different from those of either of the other two groups. However, as a group they did show significant changes in responding from pretest to posttest, indicating that acquisition of fear had occurred, although in somewhat attenuated form. (See Mineka & Cook, 1986, for further details.)

One point of special interest in the results of this experiment is highlighted by our use of medians rather than means in these figures. This is because of the bimodal

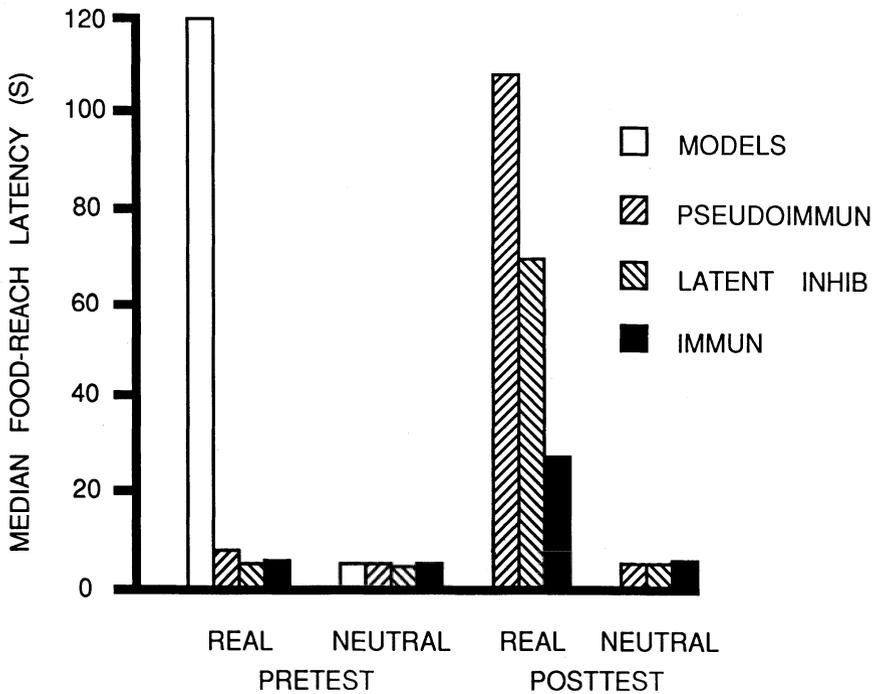


FIGURE 13. Median food-reach latency in the Wisconsin General Test Apparatus in the presence of the real snake and the neutral object for the models during the pretest, and for the Immunization (IMMUN), Pseudoimmunization (PSEUDOIMMUN), and Latent Inhibition (LATENT INHIB) groups during the pretest and posttest. (From Mineka & Cook, 1986)

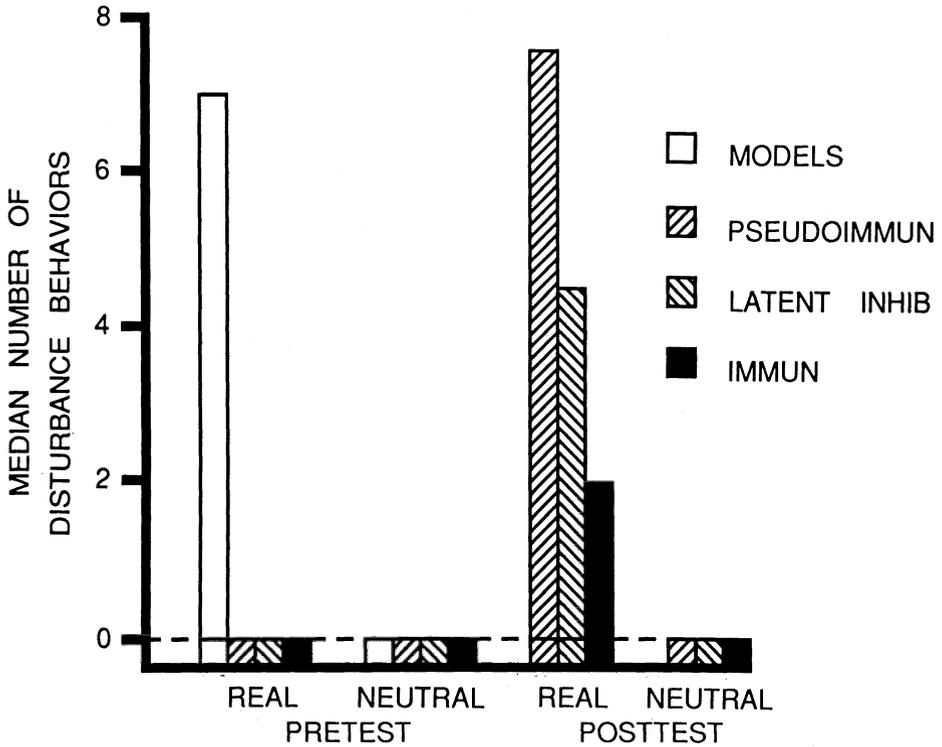


FIGURE 14. Median number of disturbance behaviors exhibited in the Wisconsin General Test Apparatus in the presence of the real snake and the neutral object for the models during the pretest, and for the Immunization (IMMUN), Pseudoimmunization (PSEUDOIMMUN), and Latent Inhibition (LATENT INHIB) groups during pretest and posttest. (From Mineka & Cook, 1986)

nature of the results in the immunization group. For six out of eight monkeys the immunization procedure was completely effective in preventing acquisition of the fear. However, the other two out of the eight showed as high levels of acquired fear as any other monkeys in the experiment. At the present time the reasons for this bimodal nature of the results remain unclear (see Mineka & Cook, 1986, for a discussion of the possibilities). Nevertheless, we do think it is quite striking that we could effectively prevent the acquisition of snake fear through simple preexposure to a nonfearful model behaving nonfearfully with snake stimuli. One might consider the results of the other observational conditioning studies described earlier to have somewhat alarming implications regarding how easily such fears may be acquired through observation alone. However, the results of the immunization study should be more reassuring that such learning can also be prevented through prior exposure to nonfearful models.

Such results may also help to account for why correlations between parental fears and children's fears are not as high as one might expect given our observational learning results. A child may sometimes have extensive preexposure to a nonfearful parent or peer behaving nonfearfully with the phobic object or situation of the other parent. Such preexposure may immunize the child against the effects of later seeing

the fearful/phobic parent behaving fearfully with that object. Indeed phobic parents may be able to prevent their children from acquiring their fears by giving their children extensive exposure to a nonfearful model interacting with their phobic object or situation.

OVERVIEW AND CONCLUSIONS

The results of the research previously described clearly implicate a strong role for social attentional processes in the origins of specific fears and phobias. Human and nonhuman primates can acquire fears not only through direct traumatic conditioning experiences, but also through vicarious experiences in which they attend to the reactions that a conspecific makes to a fear-relevant stimulus. The capacity for such social modeling obviously greatly expands the number of ways in which primates can profit from experience—not only their own experience but also that of their conspecifics. Interestingly, however, it is not at present entirely clear the degree to which the mechanisms of observational conditioning differ from those of traditional classical conditioning. In human studies of vicarious conditioning of autonomic responses, as well as in our studies of vicarious conditioning of fear, observers do not appear to be simply engaging in a social inference process. Rather, observers show significant signs of disturbance or arousal simply watching models' reacting fearfully. It is possible that this reaction to a model's fear display is like an unconditioned response to an unconditioned stimulus. (Alternatively, it may be more like a conditioned response to a conditioned stimulus if such seemingly empathetic reactions are by themselves based on a prior conditioning history.) By this view, vicarious conditioning would be mediated by essentially the same mechanisms as are first- or second-order classical conditioning.

Preliminary results on the belongingness/preparedness issue also illustrate that what is learned as a result of social attention to a model's fear display varies as a function of what the model is reacting to (e.g., snakes versus flowers), even when the fear displays are exactly equated. Such results, in conjunction with those of Öhman and his colleagues, begin to provide an account for why there is such a nonrandom distribution of the objects of people's fears and phobias. In particular, it seems likely that in the course of evolution there may have been selective pressures that enhanced the survival potential of those organisms that rapidly learned fears of certain objects or situations.

Finally, our results on immunization against the effects of observational conditioning experiences by extensive prior exposure to a nonfearful model illustrate an important point about the interactive nature of different learning experiences. In this particular case, we see the important role that everyday social attention to nonfearful conspecifics can play by interacting with and modifying the effects of later social learning experiences when other models that do react fearfully are encountered. More generally, it seems highly unlikely that most fears or phobias can be thought to originate from a single or even a few trials of classical fear conditioning or observational conditioning, occurring more or less in a vacuum, as has often been proposed in the past. Instead, there appear to be a multitude of experiential variables that can occur prior to, during, or following a traditional or an observational conditioning

experience that interact and affect the amount of fear that is experienced, conditioned, or maintained over time.

Our immunization experiment illustrates one kind of experiential variable that can occur prior to a conditioning experience that affects the amount of fear that is acquired through conditioning. Another related example of how prior experience can affect the level of fear or stress that is experienced during a frightening experience comes from an experiment done by Mineka, Gunnar, and Champoux (1986). They reared infant rhesus monkeys either in environments in which they had extensive experience with control and mastery over the receipt of a variety of reinforcers (food, water, and treats), or in environments in which the monkeys received access to the same reinforcers but not contingent on their responses. When tested between 7 and 10 months of age, the monkeys reared with control showed reduced levels of fear and higher levels of exploration in several different fear-provoking situations relative to the monkeys reared without control. Thus, early experience with control and mastery appears to affect the level of fear that a traumatic event elicits.

Mineka, Cook, and Miller (1984) have also shown that control over the offset of the unconditioned stimulus *during* the course of fear conditioning can reduce the level of fear that is conditioned to a neutral stimulus. This demonstrates that the dynamics of fear conditioning are powerfully influenced by the controllability of the unconditioned stimulus. Such results have important implications for understanding the origins of fear and anxiety disorders because many of the everyday events in which conditioning occurs are situations in which people have some control over the unconditioned stimulus.

A host of other factors occurring following acquisition of a conditioned fear response can also promote the maintenance or even exacerbation of that conditioned fear. For example, Rescorla (1974) exposed rats at a random point in time following conditioning to a higher intensity traumatic unconditioned stimulus than was involved in the original conditioning experience. The results indicated that the conditioned fear response became inflated in the direction that would be expected if the higher intensity US had been involved in the conditioning in the first place. Furthermore, Hendersen (1985) showed that the greater the time interval following the original conditioning experience and when the higher intensity US occurs, the greater the inflation effect. It is as if the organism has a memorial representation of the original US that can be altered through later experience with other USs, and the malleability of the memory increases with time. Thus a person who had a conditioning experience and acquired a fear of nonphobic intensity might be expected to show an increase in that fear, perhaps to phobic intensity if a noncontingent traumatic experience occurred at some later point in time.

These few examples serve to illustrate that the factors involved in the origins and maintenance of fears and phobias are considerably more complex than has often been assumed by behavioral learning theorists in the past. A wide range of experiential variables occurring prior to, during, or following a conditioning experience can interact and affect the level of fear that is experienced at the time and that is maintained into the future (see Mineka, 1985, for further examples). In addition, it seems certain that biological variables affect which fears are most easily acquired and maintained. Finally, as documented earlier, many fears and phobias may be acquired through

vicarious experiences watching others behave fearfully in the presence of some object, rather than through direct traumatic conditioning. Many of the same phenomena occur in the two kinds of conditioning, and indeed the mechanisms involved may not be substantially different. Nevertheless, knowledge that phobic fears can be acquired through observation alone increases awareness of how many and varied are the opportunities to acquire such fears.

ACKNOWLEDGMENTS

The author would like to thank Michael Cook for his enormous help in conducting most of the research described here.

REFERENCES

- American Psychiatric Association. (1980). *Diagnostic and statistical manual of the mental disorders* (3rd ed.). Washington, DC: Author.
- Bandura, A. (1969). *Principles of behavior modification*. New York: Holt, Rinehart, & Winston.
- Baum, M. (1970). Extinction of avoidance responding through response prevention (flooding). *Psychological Bulletin*, *74*, 276–284.
- Broadhurst, P. L. (1960). Abnormal animal behavior. In H. J. Eysenck (Ed.), *Handbook of Abnormal Psychology* (pp. 726–763). New York: Basic Books.
- Broadhurst, P. L. (1973). Animal studies bearing on abnormal behavior. In H. J. Eysenck (Ed.), *Handbook of Abnormal Psychology* (2nd ed, pp. 721–754). New York: Basic Books.
- Capitannio, J., Boccia, M., & Colaianna, D. (1985). The influence of rank on affect perception by pigtail macaques. *American Journal of Primatology*.
- Cook, E. (1983). *Human classical conditioning and the preparedness hypothesis*. Unpublished doctoral dissertation, University of Wisconsin-Madison.
- Cook, M., & Mineka, S. (1987a). Second-order conditioning and overshadowing in the observational conditional of fear in monkeys. *Behavior Research and Therapy* (in press).
- Cook, M., & Mineka, S. (1987b). Belongingness in the observational conditioning of fear in monkeys. Manuscript in preparation.
- Cook, M., Mineka, S., Wolkenstein, B., & Laitsch, K. (1985). Observational conditioning of snake fear in unrelated rhesus monkeys. *Journal of Abnormal Psychology*, *94*, 591–610.
- Cook, E., Hodes, R., & Lang, P. (1986). Preparedness and phobia: Effects of stimulus content on human visceral conditioning. *Journal of Abnormal Psychology*, *95*, 195–207.
- Delprato, D. (1980). Hereditary determinants of fears and phobias. *Behavior Therapy*, *11*, 79–103.
- Emmelkamp, P. (1982). *Phobic and obsessive-compulsive disorders: Theory, research, and practice*. New York: Plenum Press.
- Green, G., & Osborne, J. (1985). Does vicarious instigation provide support for observational learning theories? A critical review. *Psychological Bulletin*, *97*, 3–17.
- Grey, S., Sartory, G., & Rachman, S. (1979). Synchronous and desynchronous changes during fear reduction. *Behaviour Research and Therapy*, *17*, 137–148.
- Harlow, H. (1949). The formation of learning sets. *Psychological Review*, *56*, 51–65.
- Hendersen, R. (1985). Fearful memories: The motivational significance of forgetting. In F. R. Brush & J. B. Overmier (Eds.), *Affect, conditioning and cognition: Essays on the determinants of behavior* (pp. 43–53). Hillsdale, NJ: Erlbaum.
- Hodgson, R., & Rachman, S. (1974). Desynchrony in measures of fear. *Behaviour Research and Therapy*, *12*, 319–326.
- Hugdahl, K., & Kärker, A.-C. (1981). Biological vs. experiential factors in phobic conditioning. *Behaviour Research and Therapy*, *19*, 109–115.

- Joslin, J., Fletcher, H., & Emlen, J. (1964). A comparison of the responses to snakes of lab- and wild-reared rhesus monkeys. *Animal Behavior*, *12*, 348–352.
- Lang, P. (1968). Fear reduction and fear behavior: Problems in treating a construct. In J. Shlein (Ed.), *Research in psychotherapy*, *3*, 90–103.
- Lang, P. (1971). The application of psychophysiological methods to the study of psychotherapy and behavior modification. In A. Bergin & S. Garfield (Eds.), *Handbook of psychotherapy and behavior change: An Empirical Analysis* (pp. 75–125). New York: Wiley.
- Lang, P. J. (1985). The cognitive psychophysiology of emotion: Fear and anxiety. In A. Tuma & J. Maser (Eds.), *Anxiety and anxiety disorders* (pp. 131–170). Hillsdale, NJ: Erlbaum.
- Mackintosh, N. (1974). *The psychology of animal learning*. London: Academic Press.
- Mackintosh, N. (1983). *Conditioning and associative learning*. New York: Oxford University Press.
- Marks, I. (1969). *Fears and phobias*. New York: Academic Press.
- Marks, I. (1977). Phobias and obsessions: Clinical phenomena in search of a laboratory model. In J. Maser & M. Seligman (Eds.), *Psychopathology: Experimental models* (pp. 174–213). San Francisco: Freeman.
- Marks, I. (1987). *Fears, phobias and rituals*. Oxford University Press.
- McKinney, W. (1974). Animal models in psychiatry. *Perspectives in Biology and Medicine*, *17*, 529–541.
- Mineka, S. (1979). The role of fear in theories of avoidance, flooding, and extinction. *Psychological Bulletin*, *86*, 985–1010.
- Mineka, S. (1982). Depression and helplessness in primates. In H. Fitzgerald, J. Mullins, & P. Gage (Eds.), *Primate behavior and child nurturance* (Vol. 3, pp. 197–242). *The Child Nurturance Series*. New York: Plenum Press.
- Mineka, S. (1985). Animal models of anxiety-based disorders: Their usefulness and limitations. In A. Tuma & J. Maser (Eds.), *Anxiety and anxiety disorders* (pp. 199–244). Hillsdale, NJ: Erlbaum.
- Mineka, S., & Keir, R. (1983). The effects of flooding on reducing snake fear in rhesus monkeys: 6-month follow-up and further flooding. *Behaviour Research and Therapy*, *21*, 527–535.
- Mineka, S., & Kihlstrom, J. (1978). Unpredictable and uncontrollable aversive events. *Journal of Abnormal Psychology*, *87*, 256–271.
- Mineka, S., & Cook, M. (1986). Immunization against the observational conditioning of snake fear in rhesus monkeys. *Journal of Abnormal Psychology*, *95*, 307–318.
- Mineka, S., & Cook, M. (1987). Mechanisms underlying observational conditioning of fear. Manuscript in preparation.
- Mineka, S., Keir, R., & Price, V. (1980). Fear of snakes in wild- and lab-reared rhesus monkeys. *Animal Learning and Behavior*, *8*, 653–663.
- Mineka, S., Cook, M., & Miller, S. (1984). Fear conditioned with escapable and inescapable shock: The effects of a feedback stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 307–323.
- Mineka, S., Davidson, M., Cook, M., & Keir, R. (1984). Observational conditioning of snake fear in rhesus monkeys. *Journal of Abnormal Psychology*, *93*, 355–372.
- Mineka, S., Gunnar, M., & Champoux, M. (1986). Control and early socioemotional development: Infant rhesus monkeys reared in controllable versus uncontrollable environments. *Child Development*, *57*, 1241–1256.
- Murray, E., & Foote, F. (1979). The origins of fear of snakes. *Behaviour Research and Therapy*, *17*, 489–493.
- Murray, S., & King, J. (1973). Snake avoidance in feral and laboratory-reared squirrel monkeys. *Behaviour*, *47*, 281–289.
- Öhman, A. (1986). Face the beast and fear the face: Animal and social fears as prototypes for evolutionary analyses of emotion. *Psychophysiology*, *23*, 123–145.
- Öhman, A., Fredrikson, M., Hugdahl, K., & Rimino, P.-A. (1976). The premise of equipotentiality in human classical conditioning: Conditioned electrodermal responses to potentially phobic stimuli. *Journal of Experimental Psychology: General*, *105*, 313–337.
- Öhman, A., Fredrikson, M., & Hugdahl, K. (1978). Toward an experimental model for simple phobic reactions. *Behavioural Analysis and Modification*, *2*, 97–114.
- Öhman, A., Dimberg, U., & Öst, L.-G. (1985). Animal and social phobias: Biological constraints on learned fear responses. In S. Reiss & R. Bootzin (Eds.), *Theoretical issues in behavior therapy* (pp. 123–175). New York: Academic Press.

- Ost, L. G., & Hugdahl, K. (1981). Acquisition of phobias and anxiety response patterns in clinical patients. *Behaviour Research and Therapy*, *19*, 439–447.
- Pavlov, I. (1927). *Conditioned reflexes*. London: Oxford University Press.
- Rachman, S. (1977). *Fear and courage*. San Francisco: Freeman.
- Rachman, S. (1979). The return of fear. *Behaviour Research and Therapy*, *17*, 164–166.
- Rachman, S., & Hodgson, R. (1974). I. Synchrony and desynchrony in fear and avoidance. *Behaviour Research and Therapy*, *12*, 311–318.
- Rescorla, R. (1974). Effect of inflation of the unconditioned stimulus value following conditioning. *Journal of Comparative and Physiological Psychology*, *86*, 101–106.
- Rimm, D., Janda, L., Lancaster, D., Nahl, M., & Dittmar, K. (1977). An exploratory investigation of the origin and maintenance of phobias. *Behaviour Research and Therapy*, *15*, 231–238.
- Schiller, P. (1952). Innate constituents of complex responses in primates. *Psychological Review*, *59*, 177–191.
- Schwartz, B. (1974). On going back to nature: A review of Seligman and Hager's *Biological boundaries of learning*. *Journal of the Experimental Analysis of Behavior*, *21*, 183–198.
- Seligman, M. (1971). Phobias and preparedness. *Behavior Therapy*, *2*, 307–320.
- Seligman, M. E. P. (1974). Depression and learned helplessness. In R. J. Friedman & M. M. Katz (Eds.), *The psychology of depression: Contemporary theory and research* (pp. 83–113). San Francisco: W. H. Freeman.
- Seligman, M. (1975). *Helplessness: On depression, death, and development*. San Francisco: Freeman.
- Solomon, R., Kamin, L. & Wynne, L. (1953). Traumatic avoidance learning: The outcomes of several extinction procedures with dogs. *Journal of Abnormal and Social Psychology*, *48*, 291–302.
- Watson, J., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology*, *3*, 1–14.
- Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.

CHAPTER 5

Evaluative Conditioning A Case for Hedonic Transfer

A. B. Levey and Irene Martin

Our lives are governed by our preferences; likes and dislikes that we partly learn. This chapter is about one particular way in which preferences may be acquired or modified and, ultimately, about how modification of preference may be involved in therapy. The activities of daily living, when they are not merely habitual or closely constrained by the environment, obviously involve decisions and choices based on our personal preferences. More importantly, career decisions, selection of marital partners, use of spare time, and a host of other long-term choices are governed by likes and dislikes of which we may not be aware, which may be arbitrary in the sense that they have no necessary foundation, or irrational in that they contradict our best interests.

Behavior therapy, in its many ramifications, is concerned with the modification of maladaptive behavior. Most schools of practice have in common the assumption that maladaptive behaviors are acquired through some form of learning. Whether the term *learning* refers to classical conditioning, operant shaping, or reasoned argument, the assumption is that behaviors and attitudes that are acquired through learning can be modified by learning. In the light of this, it is surprising that relatively little attention has been paid to the acquisition and modification of preferences. In this chapter a case is presented for the examination of preferences both in terms of the theoretical framework of behavior therapy and in terms of possible applications.

First a basic laboratory experiment is described on which the formulation is grounded and the parameters that govern it will be explored. Then the theoretical interpretations that may be placed on the data are discussed together with reasons for believing that the phenomena described are concerned with mechanisms of classical conditioning. Lastly, some current clinical practices are surveyed and ways are suggested in which the observation and modification of preference might be used to

A. B. LEVEY • MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, England.
IRENE MARTIN • Department of Psychology, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, England.

therapeutic ends. It should be made clear from the outset that the discussion is not about complex evaluations or elaborate synthetic (e.g., esthetic) judgments. It is about simple likes and dislikes.

CLASSICAL CONDITIONING OF PREFERENCES

AN EARLY PROTOTYPE

In 1938 and 1940, Gregory Razran reported a series of experiments whose historical significance has been completely overlooked. The basic experiment was as follows. University undergraduates were asked to rate a variety of stimulus materials, photographs, paintings, literary quotations, musical selections, in terms of a set of dimensions that included personal approval, liking, social effectiveness, and so on. The stimulus materials were then divided randomly into two sets. For one set the students were allowed to select their favorite lunchtime treat, for example, a flavor of milk shake, and were provided with a free lunch during which they were continuously exposed to the stimuli. No effort was made to draw their attention to these stimuli and the setting is described as though it were informal. The other set were similarly presented while the students were required to inhale a number of unpleasant odors of a "putrid nature." From five to eight of such sessions were run for each set. Previously unrated stimulus materials were then added, in order to reduce the likelihood of remembering specific items, and the original rating procedure was repeated. Those stimuli associated with the pleasant lunch clearly showed changes in the various ratings in a positive direction; those that had been associated with unpleasant odors showed changes in a negative direction. These changes were not associated with explicit memory of the stimuli. The subjects were able to remember neither which of the stimuli were associated with pleasant or unpleasant stimuli, above a chance level, nor which had originally been seen.

It is easy to see in retrospect why these should be regarded as historic experiments. They raised two important issues, one of which went unrecognized at the time, whereas the other was of no interest to theorists of the day. Razran regarded his experiments as demonstrating a form of conditioning and the notion that human subjects would condition *outside awareness* offered no difficulties either to him or to his contemporaries. This was the issue that was of no theoretical interest. The issue that was surprisingly ignored, in view of the then current theories of conditioning, was the fact that *no demonstrable response or reflex had become conditioned*. The theories of the day required that an adequate stimulus elicit a response or reflex that then became associated with a previously neutral stimulus. What Razran had demonstrated was a change in what he called general affectivity, monitored by shifts along a dimension of pleasantness/unpleasantness, and elicited by questioning, but otherwise not apparent to an external observer.

In later experiments (Razran, 1940, 1954) he came to regard this as a form of cognitive conditioning. He showed that neutral stimuli as well as previously evaluated stimuli could be "conditioned." In other words he demonstrated the acquisition of preference, as well as shifts in preference. The notion that a central state could be conditioned, without any apparent peripheral response, was to assume enormous

theoretical importance several decades later in the realm of animal studies of so-called silent learning. Learning that can be detected only through subsequent behavioral probes is generally held to support the inference of cognition in animals (e.g., Dickinson 1980). At the time the experiments were performed, however, their significance was not recognized. Curiously, they are probably better known to social psychologists than to conditioning theorists, because the paradigm was later adopted as a means of inducing social and racial stereotypes, in studies of prejudice.

Razran was one of the pioneers of classical conditioning in the West, in terms of his originality and dedication. Having paid this tribute, we can also say that his methodology did not meet today's standards of experimental psychology. We turn now to the description of a series of experiments that, though not as picturesque as those just described, attempt to meet the requirements of experimental method.

THE EVALUATIVE CONDITIONING PARADIGM

The basic experiment will first be described as a prototype and then a number of variants will be discussed together with the parameters that control them.

Subjects are asked to choose, without deliberation, from a set of 50 postcard pictures of unfamiliar works of art or landscape photographs, the two they like the best and the two they like the least. The experimenter then pairs these pictures with those that the subject has indicated are neutral in preference. The pairs thus formed are as follows: liked preceded by neutral (forward positive conditioning); liked followed by neutral (backward positive conditioning); disliked preceded by neutral (forward negative conditioning); disliked followed by neutral (backward negative conditioning); and a further pair, neutral followed by neutral (control condition). These five pairs are presented, under passive viewing instructions, in a three-field tachistoscope for a predetermined number of trials, using appropriate controls for order effects.

Following these paired presentations, the subjects are given the stimulus cards in scrambled order and asked to sort them in rank order of liking. Having done this they are then asked to assign a numerical rating from -100 (most possible disliking) to $+100$ (most possible liking) to each of the stimulus cards. The usual result of this procedure is that neutral pictures associated in either direction with liked pictures become more liked, whereas those associated in either direction with disliked pictures become more disliked. Finally, subjects are asked whether they have noticed changes in preference, and if so, to what they attribute them. Typically, subjects are aware that changes have occurred but are not aware of their direction. They tend to explain changes as contrast effects, that is, they believe that a neutral picture should become less liked by contrast with a preferred picture even though this is not the behavior they have demonstrated.

Details of the procedure have been described more fully elsewhere (Levey & Martin, 1975; Martin & Levey, 1978). It is important to note that in presenting the stimulus pictures, considerable care is taken to frame the instructions so that the reaction of liking and disliking is based as much as possible on immediate first impressions. It is stressed that they are to use their own subjective feelings of like and dislike and this is one of the factors that make the experiment interesting. Subjects are conditioned to their own unique, individual preferences. This experiment has

been run a number of times by ourselves or by our associates in order to explore the effects of various experimental manipulations. Table I shows the essential features of the majority of these studies, and the experimental factors examined. Rather than describe any individual experiment the following descriptions will summarize the overall results. Numbers in parenthesis refer to the experiments listed in the table.

Table I represents a set of orthogonal designs from which the following conclusions may be drawn. The number of acquisition trials is not a variable of major importance (2, 9, 15). Acquisition series of 5, 10, or 20 trials do not produce differences in level of conditioning and this is consistent with the view that evaluative conditioning would be optimally adaptive if it did not depend on repeated exposures. Several studies (1, 5, 6, 7, 9) have examined the role of CS and UCS duration. Durations from one hundred to one thousand milliseconds have been tested, and have produced conditioning. This variable is of interest because the brief durations suggest that the conditioning occurs in the absence of any detailed processing of the stimulus features. This is consistent with the concept of an immediately processed evaluative response.

In the description of the procedure, given earlier, no mention was made of the methods of matching neutral pictures to the items that are liked and disliked. Three experiments (4, 5, 6, 8) have looked at this factor. If the CS items are selected deliberately so that they resemble the UCS in content and degree of detail or some important feature, conditioning is facilitated. If the pairs are selected so as to be explicitly dissimilar on some important feature then conditioning is markedly impaired. A random assignment of unselected CS items produces adequate conditioning and the technique of matching similar items has been used in some of the experiments of Table I when facilitation of the conditioning effect would make it easier to examine the experimental variable of interest.

Postexperimental questionnaires have suggested that subject awareness of the stimulus relations and demand characteristics is not an important factor, and this is

TABLE 1. Summary of Exploratory Studies of Evaluative Conditioning

Experiment	N ^a	CS	M	Experimental factor(s)
1	10	S	A	CS/UCS duration
2	10	S	C	Number of trials
3	20	S	L	Subject sophistication
4	10	S/R	C	CS matching (S or R)
5	20	S	A	CS/UCS duration
6	20	D	A	CS/UCS duration
7	20	R	A	CS/UCS duration
8	20	S/R	A	CS matching (S or R)
9	20	R	A	CS/UCS duration/trials
10	20	R	A	Subject personality
11	20	S	A	Masking task
12	20	R	A	Activity and liking
13	10	R	A	Activity only
14	10	R	A	UCS-UCS pairing
15	20	S	A	UCS intensity/one trial

^aN = number of subjects; CS = CS presentation: random (R), similar (S), explicitly dissimilar (D); M = materials: art reproductions (A), contemporary art (C), landscape photos (L).

probably due to the presentation of several pairs in a within-subject design. When this factor of awareness was explicitly examined by introducing a masking task to conceal the purpose of the experiment (11), the conditioning level was neither diminished nor enhanced. This confirms the findings of the posttest questionnaires.

Any factor that disrupts the immediate, subjective nature of the preference rating tends to reduce the likelihood of conditioning; for example, novel, avant-garde pictures (2, 4), sophisticated cognitive judgments (3, 10).

Pairing of highly evaluated pictures, either positively or negatively, is comparable to associating two unconditioned stimuli. This pairing should not produce a conditioning effect and the evidence (14) is that it does not. Similarly, if the theoretical formulation is correct, rating of dimensions other than liking should not produce conditioning. Two experiments (12, 13) tested this assumption by having subjects rate the materials for degree of activity, and this rating dimension failed to produce conditioning. Finally, the most recent experiment (15) manipulated the factor of response strength by asking subjects to select items that were strongly or only moderately liked and disliked and tested this factor against one trial conditioning. The result of this experiment, highly consistent with contemporary conditioning theory, was that single trial conditioning occurred only with the strong UCS, whereas a UCS of moderate strength could be conditioned in five trials for the majority of subjects tested under that condition.

In general the results just summarized are quite consistent with expectations for classical conditioning. It should be noted that the experiment deals with second-order conditioning, the form of conditioning that is likely to occur in real life and that has other important characteristics. Conditioning can occur to stimuli of very moderate intensity, and this is an important feature of the experiments described earlier. It would not be surprising if human subjects came to dislike a room in which they received severe electric shocks, and this would be a form of primary evaluative conditioning. However, the interest of the experiment is precisely that it deals with responses that are not physiologically active, and that lie well within the normal range of experience. Further points of interest about second-order conditioning are that it gives rise to backward conditioning and to responses that do not extinguish with nonreinforcement, neither of which is normally observed in conditioning with strong primary stimuli. Finally, the role of stimulus similarity in enhancing conditioning is a characteristic of the second-order paradigm (Rescorla, 1980). The fact that the evaluative conditioning paradigm reflects rather well the expectations for first- and second-order classical conditioning is of some theoretical interest and the next section will discuss some of its theoretical implications.

THE THEORY OF EVALUATIVE CONDITIONING

CLASSICAL CONDITIONING CONCEPTS

The idea that preferences can be conditioned is probably implicit in the concept of classical conditioning. Western writers have tended to regard Pavlovian conditioning as a kind of sterile reflexology, forgetting or ignoring the dynamic and adaptive

emphasis that Pavlov always placed on his observations. Pavlov's view was that the mechanisms of classical conditioning enable the organism to anticipate events in its environment, to respond adaptively on that basis, and to maintain a dynamic "equilibrium with the surroundings" (Pavlov, 1941a, p. 331). This involves a continuous analysis of incoming signals, and he divided their stimulus properties into those that he termed essential and those that are nonessential (Pavlov, 1941b, pp. 38ff, 82ff). The essential properties of any stimulus are those that have physiological or adaptive consequences for the organism and these are the active components of the *unconditioned* stimulus (UCS). For example, in describing the stimulus properties of food as a reinforcer Pavlov mentioned not only its capacity to evoke salivation but also noted that the food must "suit the dog's taste"; that is, it must be positively evaluated. (Pavlov, 1955, p. 142). The nonessential properties of the stimulus include its shape, color, or texture and these are the relevant components of the *conditioned* stimulus (CS).

The original basis of Pavlovian conditioning was the observation that animals salivate when they see food at a distance. It was in order to explain this phenomenon of "psychic salivation" that the systematic studies of conditioning, so often described in introductory texts, were undertaken. Razran, who devised the luncheon technique described earlier, was a close student of Pavlov's work, and there is no doubt that his demonstration of acquired preference through conditioning was derived from his knowledge of Pavlovian theory. Although it may not be appropriate in the current climate of interest to refer to Pavlov as the ultimate authority, it is important to note that the basic concept of evaluative conditioning has early and respectable origins. What the present formulation attempts to do is to give the evaluative conditioning phenomena a consistent theoretical status, to organize our own observations and those of other workers into a biologically relevant framework, and to suggest a theoretical orientation that makes the evaluative component an essential part of conditioning. Before describing this theoretical framework, some contributory studies will be reviewed.

ANTECEDENTS OF THE EVALUATIVE PARADIGM

An interesting experimental paradigm devised by Staats and Staats (1958) involves the pairing of nonsense syllables with affectively toned words over repeated trials, resulting in the acquisition of affective tone by the previously neutral nonsense syllables. The investigators regarded their work as demonstrating the conditioning of semantic meaning and the notion of verbal conditioning in this dimension has been elaborated into a systematic theoretical structure of some importance (Staats, 1975). This work will not be described further here as it is treated in the chapter by Eifert (see Chap. 8).

From time to time, investigators have become specifically interested in the preference aspects of conditioning and learning and have produced experimental results that can be seen as examples of evaluative conditioning. For example, Nunnally and her co-workers have studied the development of preferences in children and have shown that these can be modified within a conditioning paradigm. An early

experiment (Nunnally, Duchnowski, & Parker, 1965) used an ingenious game situation to induce preference for nonsense trigrams associated with three levels of reward in the form of coins: gain of two pennies for a positive outcome, loss of one penny for a negative outcome, and zero gain for a neutral outcome. The game took the form of a roulette spinwheel in which rewards looked as though they were determined by chance but were actually related systematically to the trigram stimuli. Several ingenious methods of measurement were used; for example, children were asked to pair the syllables with criterion items arranged in triplets, (sweet-bland-bitter) and to indicate which of these words reminded them most of the test trigram. Cognitive expectancy was estimated by posing hypothetical questions relating to stick figures labeled with the experimental trigrams. It is difficult to do justice to the elegance and complexity of this experiment and we are chiefly concerned here with the results. Pairing of nonsense trigrams with reward resulted in shift in the hedonic tone of the trigram toward a pleasant hedonic valence, whereas pairing with a negative outcome had a negative effect. This also applied to positive and negative cognitive expectancies.

Subsequent work by this group has shown that the conditioned reward values can be used as tokens in second-order conditioning and has demonstrated that the effects are relatively enduring. A variation of the technique has been applied to young adults (Faw & Parker, 1972) using as positive and negative stimuli the viewing of an attractive nude as opposed to immersion of the arm in ice water. These investigators showed that the effect of imaginal conditioning can be obtained with this paradigm in that merely anticipating the reward or punishment can produce conditioned alterations of the affective tone of the neutral stimuli.

Working with animals, Wyrwicka (1975) has described a number of experimental analyses that show the importance of sensory properties, including hedonic properties, in the conditioned and unconditioned stimuli for the elaboration of conditioned responses. Her description of conditioning involves the idea of "better being" and this refers to the animals' propensity to optimize desirable sensations and minimize undesirable sensations in the interests of a subjective outcome of optimum comfort. This work is in the Pavlovian tradition and extends the dynamic aspects of Pavlovian theory into the domain of evaluation.

Animal learning studies also show direct evidence in support of evaluative conditioning. Two directions of investigation are particularly relevant. A lively controversy has grown up around the conditioned acquisition of food preferences following the demonstration of taste aversion (Garcia, Ervin, & Koelling, 1966) in which animals learn to avoid distinctively flavored substances that have been poisoned. Among the significant features of this paradigm are the finding that very long delays can occur between the food ingestion and its noxious consequences and the fact that the effect can be obtained in one trial. For obvious reasons these findings raise interesting problems for conditioning theories, and a number of investigators have been wrestling with them. The issues have been well reviewed by Bolles (1983). Interestingly, these negatively conditioned preferences are sufficiently powerful to be used effectively in applied situations; for example, to deter animal predators (Ellins & Catalano, 1980) or to protect crops from the depredations of birds (Greig-Smith, 1985) without causing damaging ecological side effects.

The work on preferences has also given rise to a line of research that interacts with the cognitive view of conditioning adopted during the past decade by a number of animal workers (Mackintosh, 1983). In order to demonstrate that something has been learned, that is, that some cognitive process has occurred inside the head, one basic maneuver is to change the hedonic value of the unconditioned stimulus after conditioning has occurred. When this is done there are changes in subsequent performance that show the effect of learning that would otherwise remain silent (e.g., Holland & Rescorla, 1975; Holland & Straub, 1979; Rescorla, 1974). Alternatively, animals can learn to like the taste of noxious substances (e.g., morphine) as a consequence of conditioning to flavor preferences (Kirk-Smith, 1983). In short, preference is intrinsically involved in conditioning.

The other main area of interest in the animal research literature having to do with the role of preference in learned behavior is the phenomenon of autoshaping. A theoretical jungle has grown up around the observation by Brown and Jenkins (1968) that animals display interest in stimuli that herald rewarding or punishing events. The original observation was that pigeons being trained in a Skinnerian schedule did not need to be "shaped" to use the manipulandum, a key, provided their attention was drawn to it. If the key was illuminated, for example, just before food presentation the birds would learn to peck at it, hence the designation *autoshaping*. In order to establish that this is an instance of some form of evaluative conditioning it is necessary to sort out the locus of the effect; whether stimulus substitution occurs and the animal "thinks" that it is pecking at food or whether the response itself has a phylogenetic association, such that it pecks at anything that attracts its attention. Obviously, the point at issue is whether the animal "likes" the key. The issues are surprisingly difficult (Williams, 1981) and it is beyond the scope of this chapter to review them.

Autoconditioning is of interest in the present context because comparable behaviors were observed very early in Pavlov's laboratory whenever animals allowed to behave freely and the phenomenon came to be known as conditioning without constraint. Animals were described as behaving "joyfully," approaching the apparatus that signaled reward, wagging their tails, and generally displaying behavioral signs of what might be called positive evaluation (Zener, 1937). The issues have been carefully reviewed in a recent collection of papers (Locurto, Terrace, & Gibbon, 1981) that make fascinating reading for anyone interested in the parallels between human and animal behavior. The purpose in referring to them here is simply to indicate that the issues of preference and evaluation are alive and well in the stronghold of conditioning theory, the animal learning laboratory. As a footnote to the complex theoretical issues, one ingenious experiment signaled the delivery of food to rats by dropping another rat into the cage to serve as a conditioned stimulus. The response of the animal subjects to this conditioned signal was to make friends with it, a clear instance of positive evaluative conditioning (Timberlake & Grant, 1975).

Naturalistic behavior of this sort has been relatively little studied in animals, but would seem to offer an interesting area for the study of conditioned evaluations and their open-ended behavioral consequences. For example, many animals bury or hide objects that are negatively evaluated by them. Pinel and Wilkie (1983) reviewed a series of studies in which objects associated with electric shock were buried by rats

even after a single exposure. What is interesting about this phenomenon of conditioned defensive burying is that it is not the act of burying that is conditioned, it is the negative evaluation, to which burying the object is the animal's adaptive rejoinder.

Before turning to a summary of the theoretical position of this chapter, it may be useful merely to list references to other studies in the human literature that are relevant to the issues of evaluative conditioning but that cannot be discussed here. Zellner, Rozin, Aron, and Kulish (1983) have demonstrated conditioned enhancement of liking, whereas Sachs and Byrne (1970) showed that negative affect associated with the induction of incompatible attitudes could be transferred to the evaluation of previously neutral geometric figures. Zanna, Kiesler, and Pilkonis (1970) paired meaningful words with onset or offset of shock. Onset of shock predictably induced negative evaluation where offset produced positive evaluation and this was dependent on the degree of physiological response. These conditioned evaluations generalized to words of similar meanings. Geer (1968) demonstrated that random pairings of tones with photographs of violent death, negatively evaluated by subjects, could be used as adequate stimuli in producing GSR conditioning in the absence of any physiologically noxious stimulation, for example, electric shock. Masters and Santrock (1976) instructed children to imagine pleasant or unpleasant events and measured their effects on the maintenance of an operant task (wheel turning). This can be regarded as an evaluative analogue of the well-known conditioned emotional reaction (CER) widely studied in the animal literature. None of the studies referred to, including the animal studies described earlier, were designed as direct tests of the evaluative conditioning paradigm. They suggest, however, that evaluations play an important role in conditioning.

A THEORETICAL FRAMEWORK

We turn now to a brief description of our theoretical formulation. Some aspects of this formulation are potentially controversial, and the reader is asked to leave these on trust, for the moment, in order to present a coherent picture. The controversial issues will be addressed in a subsequent section.

The basic formulation makes the assumption that organisms have innate preferences that are correlated with the physical characteristics of their surroundings. A simple example is that all organisms "dislike" and avoid extremes of temperature without "knowing" in any sense that they are harmful. This is true from single celled organisms to the most highly developed animals and plants. Single celled organisms are also able to avoid destructive chemicals and the rejection of chemical substances that humans describe as "bitter" is almost universal among animals. This rejection of bitter tasting substances seems to have begun in the Cambrian period with the first coelenterates (Garcia & Hankins, 1975), organisms that presumably had no sense of taste in the form that we understand it. Conditioned aversion to these bitter flavors can be obtained in many species including the lowly sea anemone, and this must surely be an early form of evaluative conditioning. In another sensory dimension the primitive sea slug, *Aplysia*, has demonstrated legendary feats of learning (Carew, Hawkins, & Kandel 1983) entirely based on the fact that it "dislikes" to have its tail pinched!

On the basis of this assumption of innate preference, either general as in the case of extremes of temperature or specific as in the case of taste preferences, an evaluative response is postulated that has the following characteristics. It is immediate in the sense that it does not require cognitive mediation, decision processes, or chains of inference. That it may become involved in these cognitive processes is entirely probable, and this raises rather complex issues. However, from the standpoint of conditioning theory the treatment of evaluation as a response assigns to it much the same status as a reflex in the Pavlovian sense. It is assumed that in infant organisms this response is a powerful behavioral determinant. For example, it has recently been shown (Steiner, 1977) that human neonates have an innate hedonic preference for the sweet flavor that will induce them to ingest their mothers' milk and that this evaluative response is present before milk has been tasted. In adult humans, the response is harder to identify and it is probably seldom experienced in its innate form. Most people however will have experienced an instantaneous attraction to another person, revulsion at the sight of a mutilated animal, instant pleasure on tasting a sauce or hearing a phrase of music. These are what is meant by evaluations; they are not to be regarded as affective responses; they are not necessarily emotional.

It is not important to ask whether one must really believe that snails "dislike" acid, bacteria "dislike" extreme heat, or climbing plants "like" the sun. For the purposes of conditioning theory it does not matter. It is very probable that in organisms having a simple structure the positive evaluative response, "like," is synonymous with approach whereas the negative evaluative response, "dislike," is synonymous with avoidance. It seems clear that at some point on the evolutionary scale avoided substances began to feel unpleasant and that at some later point the organism became conscious of this feeling. What is important is that the evaluative response is a central event, and the action to be taken once it has occurred can be left open ended. This solves one of the major problems in the application of conditioning theory to the behavior of higher organisms and in particular to ourselves as sentient human beings. The Watsonian polemic that sought to reduce the complexities of human behavior to muscle twinges and glandular squirts is not acceptable in the present era. Nevertheless, we still need an emergency mechanism, left over from our Pleistocene days and nights, that allows us to cope with our surroundings: to recognize danger quickly and to identify the sources of our delights. The evaluative response is that mechanism.

One further proposition about the evaluative response is derived initially from the armchair and indirectly from experimental evidence on second-order conditioning (Rescorla, 1980). It is proposed that the evaluative response, once it is conditioned to a neutral stimulus, cannot thereafter be extinguished through nonreinforcement, but can only be altered by counterconditioning. There are several reasons for holding this view, apart from the fact that it has interesting theoretical consequences. The armchair speculation is as follows: if the evaluative component is represented in a previously neutral stimulus, as a result of conditioning, then each time that stimulus recurs it will evoke the positive or negative evaluation in its own right. This is not an unfamiliar notion in conditioning theory. The idea that responses can be self-reinforcing has been offered as an explanation, for example, of the resistance to extinction of some fear responses (Eysenck, 1976). In anecdotal terms the ice-cream cone that became a special treat in childhood retains its lifelong positive appeal even though the synthetic substitute encountered in adulthood no longer offers the same

reinforcement. The sight of its shape, color, and texture excite a positive evaluative reaction each time it occurs in spite of frequently unrewarded presentations. The reason, we suggest, is that enough of the pleasurable sensation has been transferred to the nonessential properties of the stimulus to ensure that the unreinforced presentation required for extinction can never occur.

This formulation has obvious clinical implications that will be considered later. It could be tested by an experiment in which the following three conditions were presented: (a) pairing of one strongly evaluated stimulus with another strongly evaluated stimulus (Experiment 14, Table 1); (b) pairing of an evaluatively conditioned stimulus, either positive or negative, with a strongly evaluated stimulus of the opposite valence; (c) presentation of an unpaired evaluatively conditioned stimulus. The predictions are clearly that the first pairing should produce no result and this was the outcome of Experiment 14. The second pairing, counterconditioning, should reverse the valence of the original conditioning. An experiment similar to this has been performed in the context of the children's roulette game mentioned earlier. Parker and Rugel (1973) produced conditioning of the type described, using the same financial rewards. Two weeks after the final conditioning session, subjects were told that the previously rewarded nonsense trigram would now be associated with loss of reward and vice versa. After five sessions of this new version of the game the children had reversed their original evaluations. The third pairing, extinction, in which presentation of a conditioned evaluation is followed by nothing, raises immediate difficulties. Depending on the structure of the experiment, whether within subjects or between groups, this presentation will either produce an interfering surprise that the other stimuli had been paired and this one is not; or it will produce a spurious increase in liking as a consequence of the well-known effect of mere exposure. We have yet to devise the experiment that will crucially resolve this difficulty.

To summarize, it is proposed that organisms begin life with a repertoire of innate preferences that early in life become conditioned to the specific stimulus components of the environment in which they live. The adaptive advantage of such a mechanism is obvious, in that it would enable the organism to survive in a specific environment on the basis of fairly general innate preferences. It might be noted that unless organisms had some such mechanism, survival would be very difficult in the first instance. That is, even the simplest organism needs the kind of general approach-avoidance repertoire that involves the evaluative response. Subsequent occurrences and changes in the environment are met by second-order conditioning of new preferences and the adult organism eventually has a large and complex repertoire of likes and dislikes that are entirely learned. One has only to look at cultural preferences in food, dress, or music to see the truth of this observation.

We have elsewhere presented a stronger version of this theory that says that conditioning of the evaluative response is the necessary and sufficient essential event in conditioning (Martin & Levey, 1978). For present purposes, that is in the clinical context, this stronger version serves no particularly useful purpose. A sensible and conservative view is that conditioning phenomena are made up of many components and that the evaluative component is one strong and pervasive aspect.

The formulation just described raises some problems of a different nature that have to do with the question whether what is described is really "conditioning." It might be argued that this issue is also not important in the clinical context but we

and others would disagree. Behavior therapy began with models of conditioning drawn from the human and animal laboratory and treatments based on them continue to be effective. We have previously offered (Levey & Martin, 1983; Martin & Levey, 1985) a simple rubric by which the conditioning model and the cognitive model of behavior therapy can be integrated on a single biological axis. An appropriate behavioral analysis for individual patients should reveal whether maladaptive behaviors and/or attitudes originate in the biological emergency responses (fear, avoidance, dislike) most probably associated with conditioning mechanisms or in negative interpretations of reality (self denigration, lack of confidence, expectation of defeat) likely to arise from cognitive processing, or from some mixture of both.

This view derives from Wolpe (1983), who suggests that if a disorder is cognitive in nature and is due to a distorted attitude or belief then information and persuasion are the appropriate modification procedures. If a disorder arises from classical conditioning, for example classically conditioned anxiety, then the treatment should be based on conditioning practice. Misconception correction is not ruled out by the adoption of a conditioning model. He has further suggested that inadequate behavior analysis along these two dimensions is an important source of treatment failure (Wolpe, 1977) and has laid down guidelines for treatment (Wolpe, 1981) as well as offering clinical and experimental evidence in support of his thesis (Wolpe, Lande, McNally, & Schotte, 1985). The advantage of retaining the conditioning model is that it is backed by many decades of experimental investigations.

CLINICAL APPLICATIONS

RELEVANCE TO THERAPY

The most compelling reason for advocating therapies based on the evaluative conditioning concept is that many patients present with problems that involve inappropriate likes or dislikes. Apart from the obvious paraphilias, for example, socially unacceptable, destructive, or ego-dystonic sexual preferences, likes and dislikes play an important part as symptoms. In children, for example, inappropriate food preferences, including failure to like the foods that their mothers consider to be good for them, are offered as presenting complaints, whereas problems of obesity in adults constitute an important source of clinical practice. With regard to the latter it has been suggested that quantity of food intake is conditioned to food preference (Booth, 1979) and no one would suggest that adults overeat to acquire calories. They overeat because they like the food, an evaluative preference.

Less obvious than the explicit symptoms are those cases in which some distortion of liking can be indirectly inferred. This inference can be drawn either from the description of the presenting complaints or from the case development or from the outcome in those cases in which successful treatment involves some change in preferences. This is easy to say, as an armchair dictum. Does it have any basis in reality? In order to reassure ourselves that we were not dealing merely in fantasy we reviewed 30 consecutive case reports in the most recently available complete volumes of the journal *Behavior Therapy and Experimental Psychiatry*. We chose this journal for two reasons. Firstly, case reports are routinely included and the canons of reporting are

well maintained. Secondly, the journal leans toward behavioral treatments and it was of interest to learn whether patients presenting to therapists who favor conditioning technology would show the preference trends we expected.

A comparable series of consecutive issues of the journal *Cognitive Therapy and Research* was also examined. In this journal case reports are not routinely presented but the flavor of contemporary theory is well represented. It should come as no surprise that the cognitively oriented journal is very much concerned with the problem of what people like or dislike, most often themselves. The predominance of depression and anhedonia in the concerns of cognitive therapists makes it likely that what is seen as a satisfactory treatment outcome will necessarily involve change of preference. However, the evaluative conditioning concept is probably less relevant to cognitive restructuring at this level and the following analysis is confined to the behavioral approach.

Of the 30 case reports reviewed, all but one employed some form of behavioral manipulation as the main treatment. These included relaxation, systematic desensitisation, stress reduction programs, direct reinforcement techniques both classical and operant, the manipulation of aversive consequences, time out, delayed contingency, the differential reinforcement of incompatible behaviors and so on. Thirteen of the case reports used some form of cognitive therapy as an adjunct to behavioral treatment. The criteria here included any form of explanation, rational or corrective, restructuring, discussion with significant others, self-instruction ("self-talk"), and explicit training in ideation. Seven of the cases had also used adjunctive feedback techniques, all of which were concerned with monitoring relaxation.

Exactly half the cases could be interpreted as involving aspects of preference, either in the presentation or development of the case or in the treatment outcome. Such treatments included covert sensitization and covert reinforcement (three reports) as treatment goals involving changes in preference, because this is the predetermined effect of these treatments. Some of the case categories are immediately familiar in terms of problems presenting frequently for treatment: a child likes to watch fires and takes pleasure in lighting them; a woman dislikes rather than fears the noise of traffic and this restricts her activity; a mother complains that she cannot like her child. A young woman dislikes being seen; a retarded child is trained to enjoy toileting; a little boy learns to "like" his little sister and stops quarrelling with her as an offshoot of behavioral treatment aimed at another problem. A child is cured of thumb sucking by application of a commercially available nasty tasting substance; a child who fears animals is encouraged to hold a cat and to verbalize the statement that cats are "nice"; a child who is unwilling to take fluid, to the point of dehydration, learns to like the taste of milk and this generalizes to other fluids. An adult learns to dislike alcohol.

The choice of 30 as the number of case reports was arbitrary. These problems are the grist of the therapeutic mill and will be familiar to any clinician. It is suggested, then, that where case presentation or treatment outcome involves change of preferences, where liking or disliking are inappropriate for any reason, the possibility of a conditioned origin can be considered. However, because our experiments, and those of others, demonstrate that merely associating presentations of liked or disliked materials can alter existing preferences or create new ones it seems a promising area in which to explore the development of treatment techniques.

TREATMENT TECHNIQUES

A surprisingly high proportion of aversive and punitive techniques are currently being used as behavioral treatments. Some of these, like the overcorrection and restitution methods advocated by Foxx and Azrin (1973), are mild, but they are nevertheless reminiscent of practices advocated by authoritarian parents, teachers, and sergeants major for centuries past. It is not intended to criticize these techniques nor their proponents, merely to suggest that it would be a worthwhile endeavor to attempt to develop alternatives to explicit and implicit forms of aversion therapy. Techniques such as visual screening, facial screening, the water squirt, and so on have taken the place of electric shock aversion to a large extent, but the aura of punitive manipulation remains.

Positive practice, in which the patient, usually a child, repeatedly practices appropriate forms of the relevant behavior is a constructive technique. It is also reminiscent of writing out "I must be nice to Johnny" a hundred times in an old fashioned schoolroom, and may well be perceived as punitive by the client. Similarly, the treatment role of acceptance of responsibility, for example in a restitution technique, is laudable in its intention but lies close in content to purely disciplinary practices. The hazards of aversive treatment are too well known (e.g., Walker, 1984) to need extensive discussion. They include resentment, acquired tolerance of punishment, emotional conflict, negative attitudes toward the punishing authority, and invitation of punishment for purposes of attention getting. Of 158 explicit techniques listed in one of the recent manuals (Bellack & Hersen, 1985), approximately one third could be said to be cognitive and a surprisingly high proportion of the rest could be interpreted as aversive. Similarly, an earlier survey of current trends (Sjoden, Bates, & Dockens 1979) appears to speak for a generation of authors who missed the pioneer days of behavior therapy in which misgivings about aversive treatments first developed.

To summarize, the argument is that because the conditioning of preferences can be readily demonstrated using a classical conditioning technique that avoids any aversive manipulation, it should be possible to devise treatment techniques that take advantage of it. In saying this we suggest that evaluative conditioning is separate from reward and punishment. As described earlier, the evaluative component is not emotive, and the evidence suggests that emotion disrupts it. Anyone who has had the familiar experience of watching a child throw a tantrum because he or she dislikes a particular food and refuses to eat it will be aware that the emotive behavior is not itself a function of liking or disliking. It is disruptive in contrast to the mere preference that the child uses to emotive ends. Similarly, the evaluative conditioning paradigm is not about rewarding or punishing a behavior; rather, it is about changes in perception, in attitude, or in hedonic valence.

Scattered through the literature are a few instances of therapeutic techniques that seem to be close parallels of the evaluative conditioning paradigm. Hekmat and Varian (1971), using covert semantic desensitisation, showed that snake phobias can be altered by contiguous evocation of positively evaluated images. Hekmat (1972) employed two techniques, one verbal the other imagic, to pair positive concepts with the stimulus word *spider*. The verbal emission and the visualization techniques resulted

in more positive and less fearful ratings of spiders, together with an increase in the ability to approach an actual spider although this motor behavior had not been explicitly trained.

Beech, Watts, and Poole (1971) reported the successful treatment of a young pedophile by a technique identical to evaluative conditioning in which photographs of mature women were paired with those of immature girls. More recently, a variant of Razran's luncheon technique has been used (Friedin, Borakove, & Fox 1982) to induce normal fluid ingestion in a child suffering from adipsia to the point of dehydration.

An early study by Lazarus and Abramovitz (1962) treated children's phobias by a method that could be considered to be evaluative counterconditioning: patients contiguously associated two ideas or images in a technique that the authors described as "emotive imagery." The point to be made is that a functional reanalysis of these kinds of treatment techniques, in terms of the evaluative conditioning paradigm, might well lead to the development of improved procedures. The study by Masters and Santrock (1976) mentioned earlier used the self-instruction "this is fun" in addition to imagining pleasant and unpleasant events. This possibility has not been explicitly explored in evaluative conditioning but it seems very probable that the evaluative conditioning effect could be enhanced by instructing subjects to verbalize statements such as "I like it," "this is great," and so on.

Presumably, a negative valence could be similarly induced. A recent case report may illustrate this possibility (Thorbecke & Jackson, 1982). A 19-year-old retarded female presented a problem of chronic drooling. The essence of the successful treatment package was an overcorrection procedure combined with differential reinforcement of incompatible responses, followed by a phase in which self-instruction was embedded in the treatment and substituted for the therapist's activity. As part of the treatment the young woman's attendants used expressions such as the following: "Your chin is wet, it looks horrible" or "Your chin is nice and dry, it looks lovely." In the self-instruction phase the girl was encouraged to say "Yuk! it's wet" when she had been drooling, and was then instructed to say "I have been forgetting to swallow." She was also encouraged to say "I'm dry, it looks much nicer." What is interesting about this treatment is its use of the "this is fun" method of verbal emission for inducing negative and positive evaluations. This study is also interesting because it avoids any suggestion of punitive or authoritarian treatment. The patient was encouraged to participate in her own treatment and it is clear from the description that some care was taken to ensure that she enjoyed the positive evaluations that she was encouraged to verbalize. We are not suggesting at this stage that the evaluative conditioning paradigm can simply be substituted for techniques of treatment that have evolved in applied settings, rather that clinicians be willing to consider the evaluative paradigm as a source of treatment plans.

CLINICAL ASSESSMENT

The practice of making a properly documented initial assessment or behavior analysis is one of the goals of behavior therapy, and traditionally the use of a functional analysis of the relationships among variables is the basis for treatment

intervention. In such an assessment a review of the patient's assets is likely to include some listing of likes and dislikes (preferences), for example, to be used as contingent reinforcers. This information could be extended to a functional enquiry into preference orientations that could be applied directly to the analysis of the problem behavior with a view to modifying its consequences by evaluative conditioning.

An important component of this sort of functional enquiry would be the necessity of correctly identifying and discriminating the relevant likes and dislikes from other states. It is easy to confuse a description of disliking and a description of fear. In clinical practice with children for example, it is not uncommon to find that a child genuinely dislikes a teacher or fellow pupil and this can be misinterpreted as a phobic reaction to school. It seems essential to discriminate the two. Phobic behavior is involved with a signal of oncoming events, real or imagined, whereas the reaction of disliking has purely hedonic consequences. For example, not liking to drink milk is more likely to be a simple taste aversion than a fear of the consequences of radiation hazard.

Similarly, it would seem to be important to discriminate between dislike and distress. This is not just a matter of degree, and the correct identification must rest on skillful enquiry. If the evaluative response is correct then the experience of distress is not a part of it. In the context of a threatened marital relationship, for example, it may be important to discriminate pent-up anger toward some habit or attitude that is merely disliked in the partner from genuine emotional distress based on more fundamental incompatibility. Although this type of problem has usually been the province of cognitive therapists the principle involved can be extended to other behaviors in which the appropriate treatment would include modification of the original dislike by evaluative conditioning.

Such a program would require a systematic assessment of outcome and statistical models for the assessment of preference shifts in a single case are readily available. Hersen and Barlow (1976), in their well-known text, discuss a family of two-variable regression designs that are applicable to preference shifts involved with initial baseline assessment. The most important feature of evaluative preferences in an individual is that they are uniquely the individual's own. This is the strength of the experiment that we described at the beginning of this chapter and it would similarly form the basis for applications of evaluative conditioning theory to behavior and attitude change.

In summary, it is suggested that the evaluative conditioning paradigm has potential application to clinical practice within the scope generally understood by behavior therapy. This would include the evocation of positive and or negative evaluative responses by verbal and imaginal means as well as by the presentation of explicitly evaluated stimuli.

Of existing treatments, the notion of paradoxical intention, that is, of practising an action in order to face and accept or to cognitively minimize its possible consequences, might be enhanced if the procedure included actively involving feelings of liking or disliking. Similarly, the techniques of overcorrection and restitution might be increased in effectiveness if they were combined with self-instructions designed to enhance evaluative responding in the appropriate direction. What is being suggested is that there are a number of areas in which it would be worthwhile attempting to

derive effective procedures based on evaluative conditioning. The goal of this form of therapy would be to attempt by direct means what is often attempted indirectly in conventional psychotherapy, that is, a restructuring of the relative preference valences in the client's subjective world.

ACKNOWLEDGMENTS

We are most grateful to the following colleagues and associates: Mary Acock, Hilary Ashton, Johanna Beyts, John Brown, Ruth Hodges, Annette Jedwowski, John Spinks and Maja Turcan. We are also grateful for helpful correspondence and discussions with Georg Eifert.

REFERENCES

- Beech, H. R., Watts, F., & Poole, A. D. (1971). Classical conditioning of a sexual deviation. A preliminary note. *Behavior Therapy*, 2, 400–402.
- Bellack, A. S., & Hersen, N. (1985). *Dictionary of behavior therapy techniques*. Oxford: Pergamon Press.
- Bolles, R. C. (1983). A "mixed" model of taste preference. In R. Mellgren (Ed.), *Animal cognition and behaviour*. (pp. 65–82). Amsterdam: North Holland.
- Booth, D. A. (1979). Metabolism and control of feeding in man and animals. In K. Brown & S. C. Cooper (Eds.), *Chemical influences on behavior*. (pp. 79–134). London: Academic Press.
- Brown, P. I., & Jenkins, H. M. (1968). Auto-shaping of the pigeon's key peck. *Journal of the Experimental Analysis of Behavior*, 11, 1–8.
- Carew, T. J., Hawkins, R. D., & Kandel, E. R. (1983). Differential classical conditioning of a defensive withdrawal reflex in *Aplysia Californica*. *Science*, 219, 397–400.
- Dickinson, A. J. (1980). *Contemporary animal learning theory*. Cambridge: Cambridge University Press.
- Ellins, S. R., & Catalano, S. M. (1980). Field application of the conditioned taste aversion paradigm to the control of coyote predations on sheep and turkey. *Behavioral and Neurological Biology*, 29, 532–536.
- Eysenck, H. J. (1976). The learning model of neurosis—a new approach. *Behavior Research and Therapy*, 14, 251–267.
- Faw, T. T., & Parker, R. K. (1972). Acquisition and generalisation of conditioned reward value. *Psychological Reports*, 30, 167–170.
- Fox, R., & Azrin, N. H. (1973). The elimination of autistic self-stimulatory behaviour by overcorrection. *Journal of Applied Behavior Analysis*, 6, 1–14.
- Friedin, B. D., Borakove, L. S., & Fox, K. T. (1982). Treatment of an abnormal avoidance of fluid consumption. *Journal of Behavior Therapy and Experimental Psychiatry*, 13, 85–87.
- Garcia, J., Ervin, F. R., & Koelling, R. A. (1966). Learning with prolonged delay of reinforcement. *Psychonomic Science*, 5, 121–122.
- Garcia, J., & Hankins, W. G. (1975). The evolution of bitter and the acquisition of toxiphobia. In D. A. Denton & J. P. Coghlan (Eds.), *Olfaction and taste*. (pp. 39–46). London: Academic Press.
- Geer, J. H. (1968). A test of the classical conditioning model of emotion: The use of non-painful aversive stimuli as unconditioned stimuli in a conditioning procedure. *Journal of Personality and Social Psychology*, 10, 148–156.
- Greig-Smith, P. (1985). Deter birds the tasteful way. *New Scientist*, 21, 38–40.
- Hekmat, H. (1972). The role of imagination in semantic desensitization. *Behavior Therapy*, 3, 223–231.
- Hekmat, H., & Varian, D. (1971). Behavior modification through covert semantic desensitisation. *Journal of Consulting and Clinical Psychology*, 36, 248–251.
- Hersen, M., & Barlow, D. H. (1976). *Single case experimental designs* (2nd ed.). Oxford: Pergamon Press.

- Holland, P. C., & Rescorla, R. A. (1975). The effects of two ways of devaluing the UCS after first and second order appetitive conditioning. *Journal of Experimental Psychology. Animal Behavior Processes*, 1, 355–363.
- Holland, P. C., & Straub, J. J. (1979). Differential effects of two ways of devaluing the UCS after Pavlovian appetitive conditioning. *Journal of Experimental Psychology. Animal Behavior Processes*, 5, 65–78.
- Kirk-Smith, M. D. (1983). Rats learn to like the taste of morphine. *Biological Psychology*, 17, 221–231.
- Lazarus, A. A., & Abramovitz, A. (1962). The use of “emotive imagery” in the treatment of children’s phobias. *Journal of Mental Science*, 108, 191–195.
- Levey, A. B., & Martin, I. (1975). Classical conditioning of human “evaluative” responses. *Behavior Research and Therapy*, 13, 221–226.
- Levey, A. B., & Martin, I. (1983). Cognitions, evaluations and conditioning: Rules of sequence and rules of consequence. *Advances in Behavior Research and Therapy*, 4, 181–195.
- Locurto C. M., Terrace, H. S., & Gibbon, J. (Eds.). (1981). *Autoshaping and conditioning theory*. London: Academic Press.
- Mackintosh, M. J. (1983). *Conditioning and associative learning*. Oxford: Oxford University Press.
- Martin, I., & Levey, A. B. (1978). Evaluative conditioning. *Advances in Behavior Research and Therapy*, 1, 57–102.
- Martin, I., & Levey A. B. (1985). Conditioning, evaluations and cognitions: An axis of integration. *Behaviour Research and Therapy*, 23, 167–175.
- Masters J. C., & Santrock, J. W. (1976). Studies in the self-regulation of behaviour: Effects of contingent cognitive and affective events. *Developmental Psychology*, 12, 343–348.
- Nunnally, J. C., Duchnowski, A. J., & Parker, R. K. (1965). Association of neutral objects and rewards: Effects on verbal evaluation, reward expectancy and selective attention. *Journal of Personality and Social Psychology*, 1, 270–274.
- Parker, R. K., & Rugel, R. P. (1973). The conditioning and reversal of reward value. *Child Development*, 44, 666–669.
- Pavlov, I. P. (1941a). *Lectures on conditioned reflexes*. New York: International Press.
- Pavlov, I. P. (1941b). *Conditioned reflexes and psychiatry*. New York: International Publishers.
- Pavlov, I. P. (1955). *Selected works*. Moscow: Foreign Languages Publishing House.
- Pinel, J. P. J., & Wilkie, D. M. (1983). Conditioned defensive burying: a biological and cognitive approach to avoidance learning. In R. Mellgren (Ed.), *Animal cognition and behaviour*. (pp. 39–46). Amsterdam: North Holland.
- Razran, G. H. S. (1938). Conditioning away social bias by the luncheon technique. *Psychological Bulletin*, 35, 693 (Abstract).
- Razran, G. H. S. (1940). Conditioned response changes in rating and appraising socio-political slogans. *Psychological Bulletin*, 37, 481 (Abstract).
- Razran, G. H. S. (1954). The conditioned evocation of attitudes (cognitive conditioning?). *Journal of Experimental Psychology*, 48, 278–282.
- Rescorla, R. A. (1974). Effect of inflation of the UCS value following conditioning. *Journal of Comparative and Physiological Psychology*, 86, 101–106.
- Rescorla, R. A. (1980). *Pavlovian second-order conditioning: Studies in associative learning*. London: Wiley.
- Sachs, D. H., & Byrne, D. (1970). Differential conditioning of evaluative responses to neutral stimuli through association with attitude statements. *Journal of Experimental Research in Personality*, 4, 181–185.
- Sjoden, P., Bates, S., & Dockens W. S. (Eds.). (1979). *Trends in behaviour therapy*. London: Academic Press.
- Staats, A. W. (1975). *Social behaviorism*. Homewood: Dorsey Press.
- Staats, A. W., & Staats, C. K. (1958). Attitudes established by classical conditioning. *Journal of Abnormal Psychology*, 57, 37–40.
- Steiner, J. E. (1977). Facial expressions of the neonate infant indicating the hedonics of food-related chemical stimuli. In R. Weiffenbach (Ed.), *Taste and development: The genesis of sweet preference* (pp. 173–189). Bethesda, MD: Department of Health Education and Welfare.
- Timberlake, W., & Grant, D. L. (1975). Autoshaping rats to the presentation of another rat predicting food. *Science*, 190, 690–697.
- Thorbecke, P. J., & Jackson, H. J. (1982). Reducing chronic drooling in a retarded female using a multi-treatment package. *Journal of Behavior Therapy and Experimental Psychiatry*, 13, 89–93.

- Walker, S. (1984). *Learning theory and behaviour modification*. London: Methuen.
- Williams, D. R. (1981). Bi-conditional behaviour: Conditioning without restraint. In C. M. Locurto, H. S. Terrace, & J. Gibbon, (Eds.), *Autoshaping and conditioning theory*. (pp. 55–100). London: Academic Press.
- Wolpe, J. (1977). Inadequate behaviour analysis: The Achilles heel of outcome research in behaviour therapy. *Journal of Behavior Therapy and Experimental Psychiatry*, 8, 1–3.
- Wolpe, J. (1981). The dichotomy between classically conditioned and cognitively learned anxiety. *Journal of Behavior Therapy and Experimental Psychiatry*, 12, 35–42.
- Wolpe, J. (1983). *The practice of behavior therapy* (3rd ed.). Elmsford: Pergamon Press.
- Wolpe, J., Lande, S. D., McNally, R. J., & Schotte, D. (1985). Differentiation between classically conditioned and cognitively based neurotic fears: Two pilot studies. *Journal of Behavior Therapy and Experimental Psychiatry*, 16, 287–293.
- Wyrwicka, W. (1975). The sensory nature of rewards in instrumental behaviour. *Pavlovian Journal of Biological Science*, 10, 23–51.
- Zanna, M. P., Kiesler, C. A., & Pilkonis, P. A. (1970). Positive and negative attitudinal affect established by classical conditioning. *Journal of Personality and Social Psychology*, 14, 321–328.
- Zellner, D. A., Rozin, P., Aron, M., & Kulish, C. (1983). Conditioned enhancement of human's liking for flavor by pairing with sweetness. *Learning and Motivation*, 14, 338–350.
- Zener, K. (1937). The significance of behavior accompanying conditioned salivary secretion for theories of the conditioned response. *American Journal of Psychology*, 50, 384–403.

CHAPTER 6

Knowledge, Action, and Control

Irene Martin and A. B. Levey

KNOWLEDGE, ACTION, AND CONTROL

STATEMENT OF THE PROBLEM

Underlying the cognitive and the behavioral approaches to clinical practice is a theoretical issue that is being addressed by academic theorists, but whose implications for behavior therapy are fundamentally important. The issue has appeared and reappeared a number of times but has never been satisfactorily resolved. It is the issue of the extent to which conscious awareness is involved in the control of behavior, or conversely, the extent to which effective behavior is dependant on conscious awareness.

For academic psychology the issue has come into prominence because it flows naturally from the investigation of human information processing and because recent experimental results have suggested that preattentive mechanisms account for a good deal of information processing outside awareness (e.g., Dixon, 1981; Marcel, 1983). The idea is not new and most cognitive psychologists, intimately familiar with the computer, a machine that does complex information processing outside anything that could meaningfully be called awareness, are happy with it. The older texts in the history of psychology (e.g., Flugel, 1935; Postman, 1963) provide interesting accounts of its antecedents. Helmholtz advocated the notion of unconscious inference in order to explain perception. Wundt and James simply defined psychology as the study of conscious mental activity, thereby relegating other activities of the brain to the domain of physiology.

The issues were defined for clinical psychology by the followers of Freud and Watson. Each of these giants did the field a disservice. Freud clouded the issues by

IRENE MARTIN • Department of Psychology, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, England. A. B. LEVEY • MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, England.

hypostatizing (in his structural theory) a unitary, modular unconscious, the acceptance of which precludes any serious analysis of the interactions that underlie cognitive processing. Watson, by dismissing the contents of consciousness and banning introspection and self-reports from the study of behavior, produced the identical result.

In this chapter we are concerned not with consciousness itself, but with the relationship between conscious awareness and the control of behavior. The objective of behavior therapy is the modification of maladaptive attitudes and behaviors. The realization of this objective must ultimately depend on a scientific model of action, knowledge, and control. Such a model will need to question the notion that knowledge in the head and action in the world are separate components of behavior, and in particular the view that either drives the other. The problem is to define the ways in which they interact, and this is the problem of control. A clear answer to this problem would facilitate the successful induction of behavioral change in the clinical setting.

The issues are immensely complex and difficult. In this chapter we adopt a limited approach in which the definition of awareness is confined to the notion of verbalizable knowledge. The object is to make a foray into the territory, a reconnaissance that can attempt a preliminary mapping of some of its more obvious features. To do this we concentrate on two sets of experiments, one drawn from conditioning theory, the other from cognitive theory, which are described and compared.

The issues are also not recent. The first experiments that explicitly tested the relationship between verbalizable knowledge and reinforcement history were performed by Thorndike (1932) in a paradigm that he labeled as the problem of knowledge and effect. This suggests that the formulations to be offered here are not the final answer. We begin with a discussion of the role of cognition and conditioning before turning to a description of the experiments.

COGNITION AND CONDITIONING

Behavior therapists in recent years have increasingly emphasized the role of self-awareness in controlling behavior. Concepts of self-regulation, self-management, and conscious intention have entered the clinical language and are contrasted with a stimulus-automaticity view of behavior. The ability of humans to set goals and to evaluate and change their own behavior is seen as an important source of control and as a challenge to the automaticity assumption of operant and classical conditioning. But this view of conditioning is inaccurate.

In the simplest ecological niche, direct responses to particular stimuli are guaranteed by innate mechanisms; the seeming automaticity of behavior at this level, its triggers, patterns, and sequences, present few problems to a biologically oriented behavior therapy. As the niche becomes more complex, however, the organism shows an adaptive interaction with its environment in the form of habituation, conditioning, and learning. Although conditioning theories have their roots in the stimulus-response (S-R) paradigm, they have consistently emphasized the flexibility and adaptability of the conditioned response, and the automatic growth of S-R connections is a far from adequate account of conditioned response development.

Conditioned responses exhibit subtle interactions with shifts in the environment, with context, the state of the individual, and the individual's awareness of the procedures. The present question of interest is the way in which knowledge has its effects on behavior, and these can be complex effects. Does verbalizable knowledge affect the speed of initiating a response, its persistence, consistency, the ease with which it can be inhibited, its overall pattern, or topography, or effectiveness? There are many independent and semi-independent parameters of behavior, and it has become necessary to consider more than one or two simple elements of behavior in conditioning experiments.

A similarly complex analysis needs to be applied to ways in which verbalizable knowledge relates to cognitive task performance. Recent experiments have considered for example the relation between efficient action in a situation and the subject's verbal account of these actions. Discrepancies between them are considered in terms of modules of knowledge, which may include general knowledge, that is, a data base of knowledge common to all output processes, and other specific modular processes that may interact with this common data base. One question that arises in this context concerns the representational structure of knowledge; whether for example the common data base is distinct from more specific processes that can interact among themselves and on the general data base. Some of these specific processes may result in verbal outputs, some in actions, and the knowledge that is accessible through one process may fail to reveal itself through another. Hence discrepancies arise between verbal report and action, and the complex relationship between them depends on the way knowledge is used to determine performance.

VERBALIZABLE KNOWLEDGE AND AWARENESS

It is our conviction that stimulus-based and representation-based accounts are necessary for any adequate explanation of behavior. But it would be wrong to equate conditioning explanations solely with stimulus control and cognitive theory with representational knowledge structures. It would be equally inappropriate to oppose a theory of conscious control with one that emphasizes unconscious processes. Far from being a matter of simple competition between conditioning and cognitive accounts, or between conscious and unconscious processes, the approach we wish to adopt emphasizes points of integration and investigates the nature of the interaction between them rather than assuming total dissociation.

The dissociation thesis was provocatively raised again in recent years by Nisbett and Wilson's article "Telling More than We Know," which emphasized the discrepancy between subjects' reports and their actual behavior (Nisbett & Wilson, 1977). They challenged the assumption that conscious, verbal, cognitive processes result in conscious, verbalizable changes in evaluations or motive states that then mediate change in behavior. They further argued that in the few studies where data on the issue were directly available, no association was found between degree of change in verbal report and degree of behavioral change in various experimental situations. Thus, they suggested, reports of phenomenological experience or responses based on it are conscious percepts that are affected by tacit belief systems and by nonconscious processes as much as by objective information.

Analogous evidence exists in the literature on learning and performance of manual skills, where it is generally accepted that certain crucial aspects of human performance are unavailable for introspective report. We cannot adequately describe all that we can do and conversely cannot do all that we can describe. Many authors in cognitive psychology concur that there is a dissociation between an individual's performance on a given task and the explicit or reportable knowledge associated with the task (Berry & Broadbent, 1984).

In the clinical context it has long been accepted that verbal report, behavior, and psychophysiological activity tend to show dissociation or desynchrony in emotional states such as anxiety. It has never been clear how this dissociation among components could be structured; whether, for example, they illustrate separate processes or modules, and whether or how they might interact. One model in which emotions are analyzed is the associative network, and Lang (1986) has made the interesting suggestion that information is coded in memory in the form of a network of declarative knowledge that includes response information. His emphasis on response programs serves his view that affect is basically an action set, and that the action program is associatively linked with other information, a view to be contrasted with one that holds that cognitive (verbal) and physiological responses are separate domains and independent treatment targets.

Clearly, the relationship between cognitive and conditioning models of attitude and behavior has become complex. Each has contributed to the development of behavior therapy as a body of theoretical concepts and as a source of treatment resources and techniques. In considering the effects of verbalizable knowledge on performance, then, two major categories of explanation will be addressed, one deriving from conditioning and one from cognitive contexts.

EFFECTS OF VERBALIZABLE KNOWLEDGE ON CONDITIONING PERFORMANCE

A number of differentiations must be introduced here, because the evidence as far as it goes indicates that terms like awareness, conditioning, and conditioned responses are too global to be useful, and must be broken down into specific components.

First, a number of alternative forms of awareness have been recognized as important in conditioning studies. The most obvious is knowledge of stimulus relations, which includes both conditioned stimulus–unconditioned stimulus (CS–UCS) contiguity, and the contingency between CS and UCS during acquisition, a relationship that can range from 0 to 100, be random, or explicitly unpaired (cf. Furedy, Riley, & Fredrikson, 1983). Other forms of awareness include knowledge of the experimenter's expectations (demand awareness), and knowledge by the subject of his/her own responses. Most subjects are not aware of giving autonomic, interoceptive, or eyeblink responses, but they may be aware of their responding in operant procedures that involve button pressing or other voluntary tasks.

Second, the effects of awareness are manifestly different in different response systems. It has long been known that autonomic, especially electrodermal, conditioning is peculiarly susceptible to the effects of awareness. The role of awareness in eyelid, and in skeletal-motor conditioning has not however been firmly established.

Our own studies have failed to relate conditioned response (CR) frequency to knowledge of the stimulus contiguity or to awareness of responding in a simple eyelid conditioning paradigm (Frcka, Beyts, Levey, & Martin, 1983), though one recent study reported an effect of awareness when measured as periodic expectancy during a differential paradigm (Baer & Fuhrer, 1982).

Third, the nature of the response itself must be considered. It is unrealistic to speak of "the" conditioned response as if there were a single measure denoting the whole conditioning process. Complex changes in orienting and anticipatory response frequency and topography occur during acquisition trials. It is entirely conceivable that knowledge may affect one element of responding, for example latency or frequency of occurrence, without affecting another. Finally, extinction is a process that is probably more easily affected by knowledge than is acquisition: most subjects in the eyelid conditioning situation, for example, quickly become aware of the change in stimulus schedule from acquisition to extinction and inhibit their responding.

The controversy in the human conditioning literature concerning effects of awareness on responding has in the past dealt mainly with knowledge of the CS-UCS contingency, and to a lesser extent with the role of demand characteristics and awareness of the response itself. The most extreme position within the spectrum of views on awareness argues as follows: that conditioning in human subjects can only occur through the operation of higher mental processes arising from conscious hypotheses, developed during acquisition, about the relationship between CS and UCS. Once having worked this out, subjects then try to establish what the experimenter wants them to do about it; having developed a hypothesis about the experimenter's expectations, the subject responds in order to comply with experimenter's wishes. These views on conditioning and awareness have been most explicitly proposed by Brewer (1974). He makes the assumption that there is either an awareness (cognitive) effect or a conditioning effect, and that the two are mutually exclusive. This is not a useful basis for enquiry, and the studies to be discussed in this section have been motivated to analyze the nature and extent of the interaction between them, rather than adopt an either-or position.

AUTONOMIC CONDITIONING AND STIMULUS AWARENESS

A number of experiments have been designed to analyze in detail the extent of the cognitive influence on autonomic responding. Furedy *et al.* (1983) have recorded subjects' expectancy of the occurrence of the UCS, monitored continuously by a dial that registers moment-to-moment belief about the occurrence of the UCS, under different experimental contingencies. In one group the CS is always followed by the UCS; in another the relation between CS and UCS is random (the "truly random" procedure stated by Rescorla to be the only proper control for Pavlovian conditioning); in a third, the CS is negatively correlated with the UCS (the CS being inhibitory or "explicitly unpaired" in terms of Rescorla's, 1967, contingency model).

The subjective measure of contingency or expectancy is highly correlated with the actual experimental contingency. However, such awareness is not reflected in autonomic responding, which fails to discriminate between the negative and the

random procedures. Correlations between subjects' expectancies and autonomic responding showed virtually no relationship. Further evidence of dissociation is drawn from the extinction procedure, where subjects rapidly become aware of the change in the CS-UCS contingency; autonomic responses, however, fail to parallel either cognitive awareness or the extinction operation.

Furedy *et al.* conclude that whereas the cognitive system is sensitive to propositional information about the CS/UCS contingency and reacts relatively rapidly and accurately to changes in sign-significate relationships, this mode of operation is not representative of all systems. In particular, a system such as the autonomic nervous system is relatively insensitive to CS/UCS contingencies. This evidence indicates that the two systems, one cognitive and one psychophysiological, react differently to the same procedure, and suggests that they may obey different laws. Whereas the cognitive measure of expectancy is sensitive to contiguity and contingencies, the autonomic measure seems to be sensitive only to contiguity. Such data, the authors argue, necessitate the examination of both cognitive and noncognitive processes to understand the phenomena of autonomic conditioning.

THE EYELID CONDITIONING PARADIGM

The eyelid conditioning paradigm is an example of a primitive and clear-cut form of associative conditioning that is well characterized behaviorally, and has proved extremely valuable for the analysis of theoretical issues in learning.

The blink reflex belongs to a group of specific defensive, adaptive reflexes, and shows peculiarly subtle and complex behavioral interactions with external stimuli. Depending on the nature of the stimulus schedule, the conditioned eyelid response may become more frequent, larger in amplitude, and better placed to ensure UCS avoidance. Its topography develops over trials toward a variety of endpoints such as CR/UCR integration and/or efficient avoidance of the UCS. The development of the overall response shape requires that individual response components, such as latency, amplitude, rise-time, etc., interact to determine the efficiency of the response in controlling the stimulus input (Martin & Levey, 1969).

These effects are of more than academic interest. They demonstrate even at this primitive response level many of the characteristics that go into an effective behavioral interaction with the environment. Eyelid conditioning serves as a useful and relevant laboratory paradigm for examining the implications of control. It illustrates conditioning of a specific response that produces precise CRs that are specifically adapted to the UCS.

This specificity can be contrasted with the nonspecific responses evidenced in autonomic conditioning, which are more often viewed as manifestations of a conditioned emotional state or acquired drive than as reflecting any functional interaction with environmental events. Conditioned autonomic responses presumably reflect an association between a neutral CS and the emotional properties of the UCS, in which the CS comes to elicit the same arousal/emotional components as the UCS. In addition, autonomic CRs are easily generalized to nonassociative factors, such that unpaired CSs and UCSs in autonomic conditioning schedules can lead to sensitised

and pseudo-conditioned responding. The relation of autonomic arousal to instrumental responding in this context is largely unknown.

It is hardly surprising therefore if the role of awareness differs in autonomic and eyelid conditioning paradigms with their very different implications for behavioral control.

In the set of eyelid conditioning experiments to be described, the focus of awareness has been on the pattern of the reinforcement schedule. We have asked the following question: If a set of experimental instructions provides subjects with the rules that enable them to predict the reinforcement pattern in acquisition, will overt behavior be controlled or governed by this knowledge? Two experiments manipulated the degree of subjects' knowledge of the rules governing CS-UCS pairings by means of preexperimental instructions. The question was whether subjects' knowledge of this pattern would determine the occurrence or nonoccurrence of CRs on reinforced (R) trials and unreinforced (U) trials. If knowledge is a simple determinant of responding subjects would presumably respond appropriately, that is, give responses on R trials but not on U trials. The experiments employed groups who received three differing levels of information about the stimulus schedule, one being given full information, one being asked to attend to and guess the schedule, and one given no information.

Experiment 1 employed a simple schedule in which two R trials were followed by an U trial, that is, RRU, etc. The schedule of the second experiment was more complex, and followed a rule that could be stated in two propositions (a) two consecutive R trials are always followed by a U trial, and (b) a U trial is always followed by an R trial. This schedule was more complex because a single R trial could be followed by either an R or a U trial, and hence the rule was partly indeterminate.

Results of both experiments showed that receiving instructions about the pattern of reinforcement in either a simple or complex schedule does not enable subjects to control their responding accordingly. Analyses of variance on total CR frequencies showed no significant differences between groups. Further, no significant differences were obtained between responding to reinforced and unreinforced trials within the groups. CR frequency and amplitude were only marginally less on U than on R trials, suggesting that information gained on the reinforced trials is transferred unaltered to the unreinforced trials.

These results are in agreement with others in the literature by Grant and his colleagues. An early experiment examined single and double alternation schedules (that is, RURU or RRUURRUU etc.) and found that ability to describe the schedule was unrelated to response patterns (Grant, Riopelle, & Hake, 1950). Subsequently, it was shown that subjects could learn to respond differentially to reinforced trials provided that they had some periodic feedback (Hartman & Grant, 1960, 1962; Hickok & Grant, 1964). Individual differences were noted in that some people appeared unable or unwilling to use the information available.

As a further check on awareness, subjects in both Experiments 1 and 2 were subdivided on the basis of postexperimental questionnaires into two groups: those who could state the rules correctly and those who could not. The aware groups included subjects who had not been informed of the rules, but had learned them

during the acquisition series, whereas the unaware groups included instructed subjects who could no longer verbalize them. The results are shown in Figures 1 and 2.

Figure 1 shows an initially lower level of responding to the simple schedule in unaware Ss and a steady increment in responding on both R and U trials. Aware Ss by contrast, began responding at a somewhat higher level and showed a greater discrimination between R and U trials. Figure 2 illustrates a slightly different pattern of responding under the more complex schedule. The observed differences are not statistically significant but point towards a separation of responding as a result of verbalized awareness. The evidence from both experiments supports the conclusions that (a) preexperimental instructions are insufficient to produce differential responding on R and U trials and (b) separating subjects postexperimentally according to knowledge gained by experience of the schedule shows an increased though not statistically significant ability to respond differentially.

Such results concur with others in the literature. Prokasy, Carlton, and Higgins (1967) also found that the ability to report the reinforcement pattern was not significantly related to differential performance. But they also noted other features: (a) there are marked individual differences in the tendency to use information (b) to do so, the information must be gained or "earned" by a period of exposure to the stimulus sequence; (c) elements of feedback are important for subjects to be able to utilize stimulus information.

Turning now to topographical features of conditioned responding, it has already been mentioned that response amplitudes were only marginally less on U than on R trials. This suggests that information gained on reinforced (i.e., CS + UCS trials)

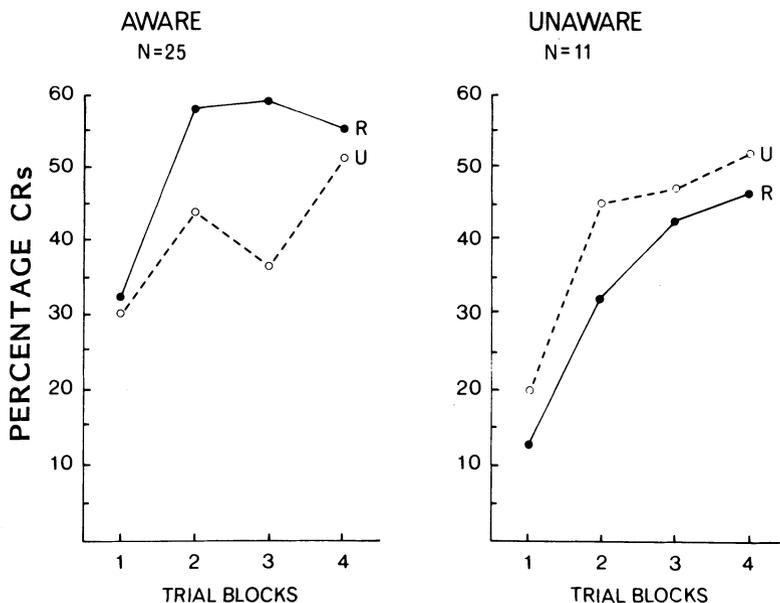


FIGURE 1. Conditioning performance of subjects divided according to reported awareness of a simple stimulus schedule: RRU, RRU, etc. R = reinforced; U = Unreinforced trials averaged separately within blocks.

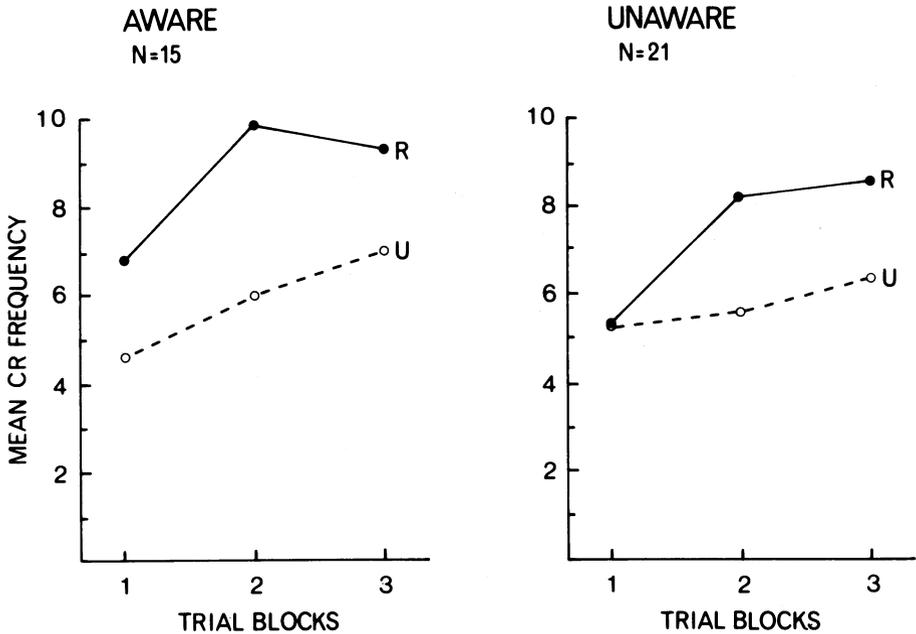


FIGURE 2. Conditioning performance of subjects divided according to reported awareness of a complex stimulus schedule. R = reinforced; U = unreinforced trials averaged separately within blocks.

is used inappropriately on unreinforced (CS alone) trials. We can infer that this process occurs outside of awareness because subjects are unable to report accurately on the frequency or amplitude of their own responses (Frcka *et al.*, 1983). This is consistent with the transfer of topographical information from reinforced to unreinforced trials through information processing mechanisms that lie outside conscious awareness (Levey & Martin, 1974). In summary, we conclude that any model of human conditioning that at one extreme attributes all conditioning performance to verbalized expectancy is incorrect. Equally compelling, however, is the conclusion that a model that at the other extreme attributes human conditioning performance entirely to trial-by-trial increments of associative strength is no more adequate.

A subsequent eyelid conditioning experiment* pursued the issue further using a different design in which a conditional discriminative stimulus, presented immediately prior to each CS-UCS pairing, contained a rule predicting whether or not the ensuing tone CS would be reinforced (Kayata, 1987). This was a row of 4 colored lights of 6-sec duration, the pattern of which predicted reinforcement or nonreinforcement: reinforcement (CS tone + UCS airpuff) could only occur if the two outer lights were the same color. No other light combination predicted reinforcement. In this design the stimuli serve as context discriminators and also as feedback insofar as subjects can check their hypotheses on each trial. The onset of colored lights in a dark room also ensured subjects' attention to them.

Four groups were used in this study. One was given no information about the rule, one was told that they could predict it, and one was fully informed with several

*We are grateful to Lisa Kayata for permission to quote details of this experiment.

demonstrations of the colored light combination to ensure understanding. In addition to these three levels of information a fourth group was misinformed. Electrodermal responding was monitored throughout.

The results of this experiment showed clear and highly significant differences in conditioned responding in the fully informed group (see Figure 3). No other group showed this significant discrimination. The fully informed group also showed significant discrimination of electrodermal responding to the onset of the light stimuli. Thus in this new paradigm instructions are shown to affect conditioning performance and concomitant electrodermal activity.

All subjects in the fully instructed group could reproduce the rule governing reinforcement both preexperimentally and postexperimentally; in addition some subjects from the other uninstructed groups had guessed the rule correctly through exposure to the conditioning sequence and could verbalize it postexperimentally. These "became aware" subjects were compared with the fully instructed subjects, and the results are illustrated in Figure 4. Comparing this figure with the previous one, the main difference between them is that the separation on initial trials is greatly reduced, indicating that the information necessary to achieve differentiation was acquired during these early trials. Clearly, when subjects are reminded on each trial

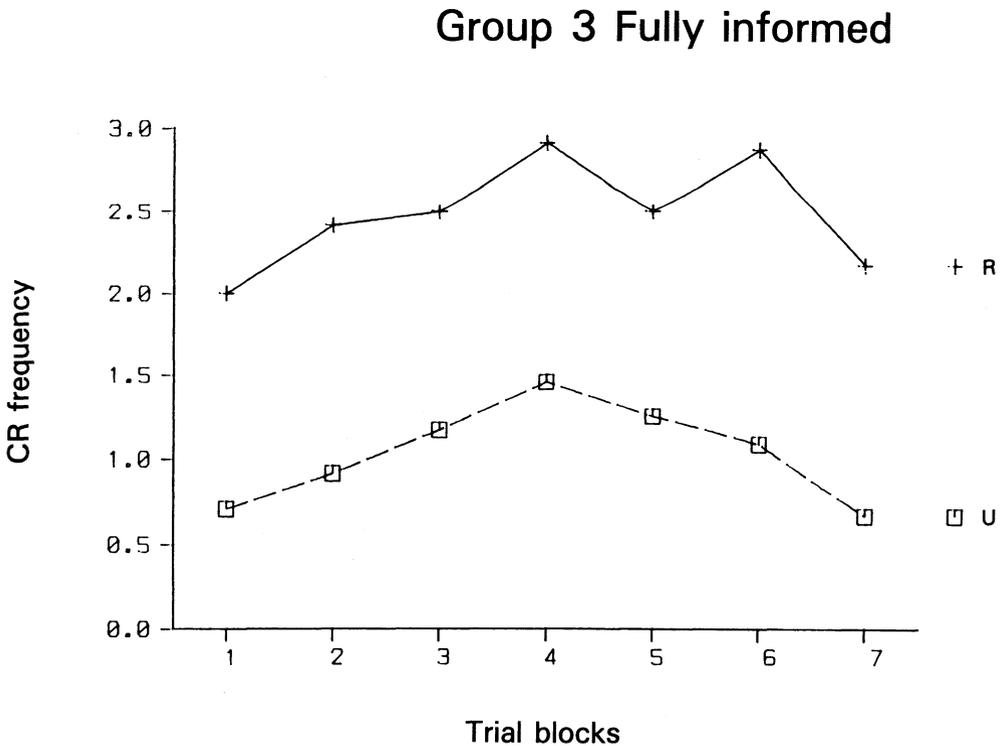


FIGURE 3. Conditioning performance of subjects given full preexperimental instructions of stimulus schedule plus a trial-by-trial reminder. R = reinforced, U = unreinforced trials averaged separately within blocks.

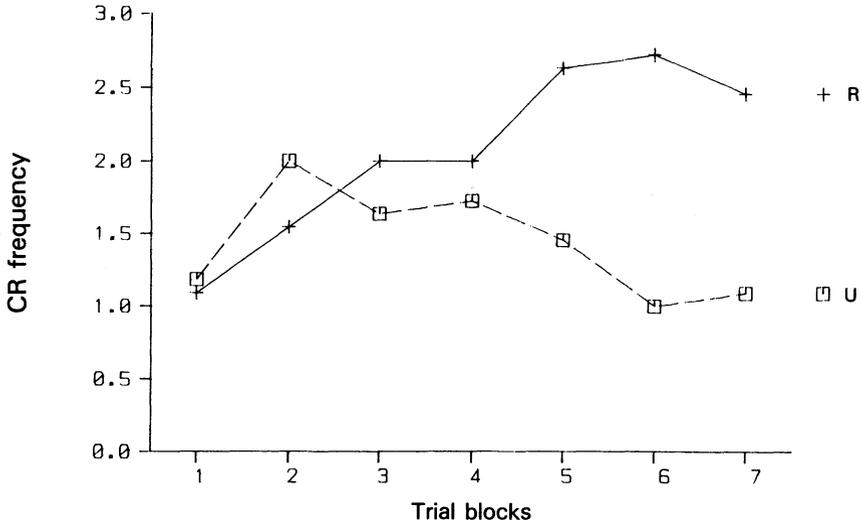


FIGURE 4. Conditioning performance of subjects divided according to post-experimental report of awareness. R = reinforced, U = unreinforced trials averaged separately within blocks. N = 11.

of the verbalizable rule governing reinforcement the growth of awareness contributes significantly to their conditioning performance.

These three experiments illustrate the attempt to examine effects of awareness when the conditioning paradigm involves a rule which governs the occurrence of reinforcement. Taken together, the results shown in these figures demonstrate that subjects can use information to modulate responding when given preexperimental instructions and also when they acquire knowledge through experience. The main feature of significance that contributes to this finding seems to be the 6-sec light stimulus preceding each CS-UCS trial in experiment 3, which can serve as a predictor of what is about to happen. The next section compares these results with performance on a cognitive task.

EFFECTS OF VERBALIZABLE KNOWLEDGE ON COGNITIVE PERFORMANCE

The research considered in this section derives from current work in cognitive psychology, and has been selected for discussion because it focuses specifically on the relationship between verbal report and performance. The studies are of particular relevance in that, like the conditioning paradigm, the tasks involve practice over trials and so permit assessment of changes in performance as well as change in verbal report.

A recent study of this type by Berry and Broadbent (1984) reported results that are consonant with those of the experiments just described. Observing subjects' ability to use a simple algorithm (analogous to the metaknowledge of the summary rules in the conditioning experiments) to control a cybernetic production system,

they found that performance improved with practice but not with instruction, and that verbal report of strategies bore little relation to levels of performance success. They also noted individual differences in the ability to use information, and found that trial-by-trial correction of verbalized strategies assisted learning where mere verbalization did not.

INFORMATION LOAD AND SALIENCE

These and similar findings raise several questions that are tackled in a subsequent series of experiments (Broadbent, Fitzgerald, & Broadbent, 1986). The first problem was to establish the generality of the observation that verbal report and performance are dissociated; whether, for example, it is a function of the complexity of the information load. If the individual is required to learn a very simple relationship, for example, that A always leads to B, she or he is likely to utilize such information in performance. If the task involves knowledge of a more complex relationship between two variables, knowledge that could be consistent with a large number of possible decisions being correct, the information load is substantially increased. Broadbent *et al.* used as their task a model economic system having two parameters or rules: increasing government expenditure decreases unemployment and increasing taxation increases unemployment. Subjects were given initial values of these two parameters and asked to manipulate them to achieve a self-determined target of unemployment. The nicety of this model is that even knowing these two rules, there is still a large number of values of taxation and expenditure consistent with a desired level of unemployment. "Thus it is very plausible that a person might possess verbal knowledge about economics and yet be unable to take correct decisions, just as they may be able to take correct decisions and yet be unable to answer verbal questions."

One of the experimental manipulations was therefore level of information load. They also included a factor of salience, represented by the size of change produced by a decision. For one group, the excessive consequences of mistaken decisions were corrected by the computer; for another this feature was removed with the result that the effects of mistaken decisions persisted, presumably with dramatic impact on the subject's awareness. It was predicted that such a system would produce better verbal knowledge but less effective performance.

Results were in line with the authors' expectations: quantity of information and saliency of presentation related to verbal knowledge. Good verbal knowledge depends on having relatively few variables in the situation and making the key variables salient. However, increased salience failed to show a corresponding increase in the effects of practice on performance. Overall, practice with the tasks increased the probability of correct performance, without a corresponding increase in the number of correct answers to verbal questions.

An important point from the theoretical perspective of these investigators is that the relationship between verbal report and performance is not asymmetric: there can be cases of good verbal knowledge with poor performance as well as the reverse. That dissociation between verbal knowledge and performance could be changed by altering the number or salience of the relationships being learned, and the fact that

one but not the other could change with practice, suggests that the experimental variables affect specific information processes rather than change a common data base.

Again, these general results are consistent with older findings. Postman and Jarrett (1952), investigating a card-calling experiment in which subjects were required to learn the significance of hidden cues, reported the following conclusions: (a) improvement in performance over a series of trials precedes correct verbalization; (b) the ability to verbalize improves performance; but (c) the ability to verbalize is not necessary to correct performance and does not lead to error free performance.

ALTERNATIVE THEORETICAL MODELS

The explanation put forward by Broadbent *et al.* is phrased in terms of modules of knowledge. A commonsense view, which the authors question, supposes that people act by consulting an internal model of the world, a data base of knowledge common to all output processes, and manipulate it to decide on the best action. To handle discrepancies between verbal report and action, this view must also suppose a distinction between the general data base of knowledge and other relatively specific processes that act on it, such that some of the outputs will result in verbal outputs, and some in actions. A similar issue has recently been discussed by Fodor (1985), who captures the spirit of this enquiry by proposing a view of cognitive architecture that emphasizes the distinctions rather than the continuity of modular processes. He argues that certain processes (perceptual processes in his example) are encapsulated from the general background data base of knowledge. The problem lies in specifying their relationship to one another: whether the modules are interactive and permeable or encapsulated and isolated.

The experiments just described raise an additional issue, the need for some kind of mechanism to explain where behavior comes from. Cognitive psychology has long been criticized for its neglect of behavioral determinants and its assumption that if only we could understand what the organism knows and how it organizes its thoughts about the environment its behavior would fall out simply and automatically. The authors are aware of this problem and discuss some of the alternative possibilities. If, as their data suggest, individuals do not consult a common data base and act on decisions arising from such consultation—the model-manipulation view—what alternatives can be considered? One possibility is a situation-matching model in which the individual identifies key features of the situation and decides either on the basis of those alone or on the basis of similarity between the situation now present and others encountered in the past. A highly similar sequence from the past will resemble the present one in many irrelevant features but it will have a greater than chance probability of resembling it in the key ones as well. Thus action based on matching the current situation to one from the past will give better than chance performance although it might be incapable of providing verbal answers to questions about performance.

The authors consider the inadequacies of such a model: it simply suggests that people can sometimes make the right decisions although they do not know the rules. If they can state the rule verbally they are more likely to identify the key features of

the situation and decide on the basis of those alone. However, restricting the argument to the case where the individual cannot verbalize rules, the problem remains of how decisions are made that lead to action. Two extreme strategies are described for purposes of illustration, one of which calculates the future outcome of each possible action, using observation of the current situation and knowledge of the structure of the world: the look-ahead strategy. The alternative is one that stores a previously generated table that records the correct action to be taken in each of a variety of situations. This, the situation-matching strategy, is termed the look-up table alternative. The distinction between them highlights the role of knowledge in behavior. Whereas the look-up table system can give no account of the reasons for the particular action it chooses, the look-ahead system can justify its actions by comparing their expected effects with those of other actions, because the latter have also been calculated.

Although the issues raised by these formulations seem more complex than those raised by the conditioning results, this is partly a matter of differing styles of investigation. Comparing these two groups of studies, one drawn from the conditioning model of behavior, the other from contemporary cognitive theory, the data they yield have considerably more in common than the conceptual backgrounds that motivate them. This suggests that the two approaches overlap in this area and may be profitably combined. In the next section we consider some of the obstacles and encouragements to reach a combination of interests.

INTEGRATION OF BEHAVIOR STIMULUS CONTROL AND MENTAL REPRESENTATIONS

One of the major dissatisfactions with conditioning explanations of behavior centers on the issue of stimulus control—the implication of individuals as robots, being passively conditioned, mechanically controlled by events in the environment. By contrast, if knowledge is construed as a complex mental structure, internal rather than external, control seems to be handed back to the individual in terms of self-regulation and self-management. The thesis that stimuli exert control over behavior has been widely proposed, discussed, and criticized in the past. Although there can be little doubt about the reality of stimulus-reinforcer controls on behavior there are many mysteries about the nature of these effects. One concerns the enigmatic nature of the stimulus itself, and the analysis of the stimulus into components, some of which may be salient whereas others are irrelevant. Stimulus configurations are of special interest to contemporary conditioning theories because they throw light on the ways on which stimuli are analyzed and represented.

This is illustrated in Rescorla's recent animal work on the way in which qualitative perceptual relations may affect conditioning. When a CS is paired with a reinforcer that has multiple aspects the animal selects from the reinforcer those aspects that it will encode. For example, when it has to associate two multifeatured events, it will learn about the co-occurrences within each event as well as forming associations between events. Rescorla considers what the relation might be between these two kinds of learning, suggesting that the organism integrates within each event and then associates those integrated representations with each other, and he draws an analogy with Gestalt principles of proximity, similarity, and closure (Rescorla, 1985).

It is noteworthy that this kind of analysis of stimulus components, effective elements, and the active perceptual engagement of the individual in selecting, integrating, and coding this information within some kind of representational system should be generated by conditioning theorists. It emphasizes a marked shift from viewing stimuli as external to the individual and exerting environmental control to analyzing the interaction between individual and environment, and further to suggesting ways in which the information might be stored as stimulus representations. Thus the criticism that conditioning models of human behavior ignore mediating events is unjustified. In fact no important behavior theory ever conformed to a simplistic reflexological model. From the earliest days, principles of learning, internal stimuli, integration, and coordination were included, all of which violated the terms of a reflexological model (Zuriff, 1985).

THE ARCHITECTURE OF COGNITION

If behavioral psychologists have been obsessed with describing functional relationships between stimuli and responses, cognitive theorists are preoccupied with models of mental architecture to the neglect of determinants of overt behavior. Fodor (1985) refers to the "baroque proliferation of scripts, plans, frames, schemata, special-purpose heuristics" and critically examines the varieties of architectural structures that have been proposed, in particular that of a higher-order cognitive process (thinking, awareness, problem solving) that operates in a top down fashion to influence and penetrate different specific processes. This interactionist view of cognitive processes is questioned by Fodor, who discusses an alternative that he refers to as encapsulation. The general issue is the boundary of interactions between knowledge sources or modules: how rigid the boundary is between the information available to cognitive processes and the information available to perceptual or reflexive processes, for example, or to behavior.

A number of formal systems have been developed, such as parallel distributed processing schemes, which do not necessarily assume a modular theory of mind. Such models assume that information processing takes place through the interaction of a large number of simple processing elements or units, each sending simple excitatory and inhibitory signals to one another. An important advantage over earlier information processing models is that these models can incorporate learning mechanisms and their units can include response components (Sutton & Barto, 1981). One criticism of this approach, however, is that although these formal systems are sufficient for the purpose of representing any computationally explicit model of cognition, they are less sensitive to the problems that arise in attempting to develop explanatory theories within such frameworks.

CLINICAL CONDITIONS

In the clinical context, the issue of awareness centers on those cognitive approaches to therapy that assume that remodeling the individual's verbalizable statements about himself will lead to adaptive changes in behavior. The role of

self-talk, self-instruction, conscious modeling, and so on, seems to stress the importance of verbalizable knowledge in the control of behavior. It has already been suggested that behavior therapy must eventually adopt a specifiable model of the relationship between knowledge and action to serve as a basis for predictable changes in attitude and behavior, and this relationship has been examined in both cognitive and conditioning models. Shifts in orientation in these disciplines suggest that some of the old distinctions are being eroded. Conditioning studies of animal behavior, for example, offer cognitive models that show how principles of inference allow the animal to incorporate into its stimulus representations something about the causal structure of the environment (Dickinson, 1980).

Thus some communality of interest appears to be emerging as conditioning and cognitive theorist approach such academic problems as the attribution of causality (Dickinson, Shanks, & Evenden, 1984), the interaction of stimulus components (Rescorla, 1985), and the relationship between knowledge and control (Broadbent & Berry, 1984). All of these issues are related to the central problem of verbalizable knowledge and awareness.

Whatever understanding the future development of cognitive models may offer, there are as yet few useful or relevant explanations from laboratory research that are applicable to the control or modification of overt behavior. As a result, variants of cognitive theory are being generated from practice. In an earlier era it was the behavioral approaches, allegedly grounded in learning theory, that gradually drifted toward a more pragmatic orientation in which theory was modified to suit practice. There is now a similar danger that cognitive explanations will be offered merely for the sake of adopting a fashionable point of view, just as conditioning explanations were once offered for clinical phenomena (e.g., depression) that lay outside their scope.

When behavior therapists refer for example to cognitive misinterpretation as a determinant of avoidance behavior, we have to consider what evidence would justify this as a cognitive hypothesis. Clark's (1986) cognitive model of panic attacks views them as a catastrophic misinterpretation of certain bodily sensations: people feel they are going to die, go mad, lose control. It seems possible to reinterpret this cognitive model along traditional lines, as a conditioned fear of fear, or conditioned negative evaluations of unpleasant and distressing bodily sensations. We have made a similar suggestion elsewhere concerning cognitive interpretations of the modeling phenomena, pointing out that modeled fears can be explained along conditioning lines (Levey & Martin 1983; see also Mineka, Chapter 4). Observing someone experiencing agonizing fear must rank as a most potent and frightening stimulus that can therefore function as a UCS in a conditioning paradigm.

Now that the era of useless competitiveness between conditioning and cognitive explanations is hopefully ended, it becomes important to establish real cognitive hypotheses, and to be wary of so-called cognitive interpretations that are pale versions of already existing conditioning explanations. Cognitive psychology has obviously a lot to offer, but its preoccupations and formulations have not been with behavioral determinants *per se*. The way individuals interpret what happens to them may be of utmost importance to how they react: the point is simply that we cannot specify with any clarity how a cognitive explanation would work, and how it could be distinguished

from a conditioning explanation, until we can specify how cognitive processes that lead to changes in verbalized statements interact with behavior.

The most obvious examples of dissociation between knowledge and control probably occur in the treatment of complex addictions, for example, tobacco and alcohol. Knowing that smoking is potentially dangerous is not usually helpful in changing smoking behavior. On the other hand there is ample epidemiological evidence that this knowledge has led to a marked reduction in smoking. Similarly, it is possible to condition an alcoholic to vomit on exposure to alcohol without thereby preventing further ingestion unless the patient has accepted some form of cognitively mediated will to stop.

It can be asked of maladaptive behavior, generally, whether individuals are aware of their consequences and whether this knowledge reduces the frequency of the behavior. An area in which the cognitive therapies have been notably successful, social skills training, provides an interesting example on the distinction between not knowing the social rules and not knowing how to apply them. The former is analogous to the uninformed or misinformed subjects, the latter to the growth of awareness in the experiments described earlier. The actual acquisition of skills through practice, including experience in the interpretation of expressive cues, is reminiscent of the finding in the cognitively based and conditioning experiments that subjects need to earn knowledge by responding to the stimuli and that reminders and feedback play a facilitating role in this process. It can be hoped that further studies of the exact conditions under which knowledge facilitates effective behavior will feed back into clinical practical.

CONCLUSIONS

In this chapter we have considered the relationship between knowledge and action in the control of behavior. Experimental evidence on awareness, verbalizable knowledge, and their effects on behavior has been examined in two contrasting response paradigms: eyelid conditioning and complex cognitive tasks involving the management of control systems. This analysis of two types of laboratory tasks illustrates that the very different conditioning and cognitive approaches to behavioral change as a function of learning and practice have surprisingly similar consequences in their results, in spite of the differing conceptual and theoretical positions that each adopts. It also confronts a practical issue in behavior therapy: how knowledge that is verbalized can modify behavior.

Examining the relationship between verbalizable knowledge and overt behavior, whether behavior is in the form of conditioned responses or complex decision-based performance, strongly suggests the need to integrate cognitive approaches to knowledge with conditioning approaches to behavior. Evidence showing that verbalizable knowledge can affect conditioned responding, together with the view outlined in the previous chapter that behavior and attitudes are governed by preferences, that is, likes and dislikes that are often irrational and beyond conscious access, reflect determinants of behavior change that are best considered within the recent developments of conditioning theory. Contemporary conditioning theories provide a sophisticated

structure in which to consider how evaluations can be associated with different stimuli and how individuals learn about the causal structure of their world—two very different and very important sources of learning that affect behavior.

If, however, behavior is also determined by long-term storage of experience then we need to know more about the structure of such knowledge and how it leads to behavior. One implication of integrating conditioning and cognitive approaches to behavior is to recognize that some aspects of behavior are under stimulus control whereas other aspects are guided by mental representations of knowledge. Such an implication no longer carries the conviction that all conditioned behavior is simplistic, automatic, and entirely stimulus controlled. Behavioral theory has gone a long way toward defining external determinants in its analysis of stimulus control and stimulus-response relationships, and contemporary conditioning theory is now moving inward to shift some of the external control of behavior to mediating internal representations of events.

The part of cognitive theory that is of relevance in the present context is that which attempts to suggest how the structure of knowledge interacts with decision processes to eventuate in behavior. Some part of the individual's knowledge can be verbalized, and the term *verbalizable knowledge* has been used here, as a convenient definition of awareness, to refer to the individual's ability to define a stimulus rule, either through the provision of verbal instructions by the experimenter or as a result of the subject's own verbalization to questions following exposure and practice.

Awareness, in the sense defined, is related in a complex manner to action in the world. Awareness seems to modify action but does not determine it; action seems to modify awareness by providing knowledge through experience. Jointly, these processes can lead to the understanding of control.

REFERENCES

- Baer, P. E., & Fuhrer, M. J. (1982). Cognitive factors in the concurrent differential conditioning of eyelid and skin conductance responses. *Memory and Cognition*, *10*, 135–140.
- Berry, D. C., & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalisable knowledge. *Quarterly Journal of Experimental Psychology*, *36A*, 209–231.
- Brewer, N. F. (1974). There is no convincing evidence for operant or classical conditioning in adult humans. In W. Weimer & D. Palermo (Eds.), *Cognition and the symbolic processes* (pp. 1–42). Hillsdale, NJ: Erlbaum.
- Broadbent, D. E., FitzGerald, P., & Broadbent, M. H. P. (1986). Implicit and explicit knowledge in the control of complex systems. *British Journal of Psychology*, *77*, 33–50.
- Clark, D. M. (1986). A Cognitive Approach to Panic. *Behaviour Research and Therapy*, *24*, 461–470.
- Dickinson, A. (1980). *Contemporary animal learning theory*. Cambridge: Cambridge University Press.
- Dickinson, A., Shanks, D., & Evenden, J. (1984). Judgement of act-outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology*, *36A*, 29–50.
- Dixon, N. (1981). *Preconscious processing*. New York: Wiley.
- Flugel, J. C. (1935). *A hundred years of psychology*. London: Suchworth.
- Fodor, J. A. (1985). Précis of the modularity of mind. *The Behavioral and Brain Sciences*, *8*, 1–42.
- Frcka, G., Beyts J., Levey, A. B., & Martin, I. (1983). The role of awareness in human conditioning. *Pavlovian Journal of Biological Science*, *18*, 69–76.
- Furedy, J. J., Riley, D. M., & Fredrickson, M. (1983). Pavlovian extinction, phobias, and the limits of the cognitive paradigm. *Pavlovian Journal of Biological Science*, *18*, 126–135.

- Grant, D. A., Riopelle, A. J., & Hake, H. W. (1950). Resistance to extinction and the pattern of reinforcement. I. Alternation of reinforcement and the conditioned eyelid response. *Journal of Experimental Psychology*, *40*, 53–60.
- Hartman, T. F., & Grant, D. A. (1960). Effect of intermittent reinforcement on acquisition, extinction and spontaneous recovery of the conditioned eyelid response. *Journal of Experimental Psychology*, *60*, 89–96.
- Hartman, T. F., & Grant, D. A. (1962). Effects of pattern of reinforcement and verbal information on acquisition, extinction and spontaneous recovery of the conditioned eyelid response. *Journal of Experimental Psychology*, *63*, 217–226.
- Hickok, C. W., & Grant, D. A. (1964). Effects of pattern of reinforcement and verbal information on acquisition and extinction of the conditioned eyelid response. *Journal of General Psychology*, *71*, 279–289.
- Kayata, L. (1987). Awareness and human learning. Unpublished doctoral thesis, University of London.
- Lang, P. J. (1986). The cognitive psychophysiology of emotion: Fear and anxiety. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders*. Hillsdale, NJ: Erlbaum.
- Levey, A. B., & Martin, I. (1974). Sequence of response development in human eyelid conditioning. *Journal of Experimental Psychology*, *102*, 678–686.
- Levey, A. B., & Martin, I. (1983). Cognitions, evaluations and conditioning: Rules of Sequence and Rules of Consequences. *Advances in Behaviour Research and Therapy*, *4*, 181–195.
- Marcel, A. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology*, *15*, 197–237.
- Martin, I., & Levey, A. B. (1969). *The genesis of the classical conditioned response*. Oxford: Pergamon Press.
- Nisbett, R. E., & T. D. Wilson. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review*, *84*, 231–259.
- Postman, L. (1963). (Ed.). *Psychology in the making*. New York: Knopf.
- Postman, L., & Jarrett, T. (1952). An experimental analysis of learning without awareness. *American Psychologist*, *65*, 244–255.
- Prokasy, W. F., Carlton, R. A., & Higgins, J. D. (1967). Effects of nonrandom intermittent reinforcement schedules in human eyelid conditioning. *Journal of Experimental Psychology*, *74*, 282–288.
- Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review*, *74*, 71–80.
- Rescorla, R. A. (1985). Pavlovian conditioning analogues to Gestalt perceptual principles. In F. R. Brush & J. B. Overmier (Eds.), *Affect, conditioning, and cognition: Essays on the determinants of behavior* (pp. 113–130). Hillsdale, NJ: Erlbaum.
- Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, *88*, 135–170.
- Thorndike, E. L. (1932). *The fundamentals of learning*. New York: Teachers College, Columbia University.
- Zuriff, G. E. (1985). *Behaviorism: A conceptual reconstruction*. New York: Columbia University Press.

CHAPTER 7

Operant Conditioning The Hiatus between Theory and Practice in Clinical Psychology

C. F. Lowe, P. J. Horne, and P. J. Higson

INTRODUCTION

The application of operant conditioning principles, characterized variously as behavior modification or applied behavior analysis, has become widespread in clinical psychology. A central theoretical assumption underlying this approach is that operant principles, originally derived from the study of animals in controlled experimental settings, have general applicability, governing not only the behavior of animals but also that of humans. In their early work, Skinner and others showed that animal behavior was an orderly function of contingencies of reinforcement so that any particular performance, on a schedule of reinforcement for instance, could be analyzed within the framework of the “three-term contingency,” that is, the relationship between responses, reinforcers, and discriminative stimuli. The response was usually the operation of some mechanical device like a lever; the reinforcer was typically food, and discriminative stimuli were environmental events, such as the illumination of colored lights. All of these variables were publicly observable events. The creation of explanatory fictions, “events taking place somewhere else, at some other level of observation” (Skinner, 1950), was eschewed.

This then was the model adopted by the behavior modification movement. Focused as it was on observable behavior and environmental stimuli, it was taken by many to exclude all consideration of covert or “cognitive” events. The basic conditioning principles were established, and all that remained was the development

This chapter contains some material from an earlier chapter “Is All Behavior Modification Cognitive?” by C. F. Lowe and P. J. Higson in E. Karas (Ed.), *Current Issues in Clinical Psychology*, 1983, Plenum Press, New York.

C. F. LOWE and P. J. HORNE • Department of Psychology, University College of North Wales, Bangor, Gwynedd, LL57 2DG, Wales. P. J. HIGSON • Department of Psychology, North Wales Hospital, Denbigh, Clwyd LL16 5SS, Wales

of an appropriate technology for their application to clinical and other social problems.

In recent years, however, it has become apparent that in spite of this early confidence the power to predict and control complex human behavior has proved elusive. Indeed, behavior modification's manifest lack of efficacy in some areas has led to a good deal of soul-searching within the area (cf. Branch & Malagodi, 1980; Dietz, 1978; Hayes, Rincover, & Solnick, 1980; Michael, 1980, 1985). Simultaneously, a number of critical theoretical issues have been raised, foremost among these being the question of the relationship between behavioral theory and applied work (Epling & Pierce, 1986; Kohler & Greenwood, 1986). For, although the success of behavior modification procedures has often been taken as evidence that human behavior conforms to the same laws of learning as apply to animals, a number of authors have argued that the effectiveness of such procedures owes little to conditioning principles, and some have gone so far as to claim that there is, in fact, no convincing evidence for either operant or classical conditioning in humans (see Beech, 1974; Bloomfield, 1976; Boulgouris, 1982; Brewer, 1974; Dietz, 1978). Studies in the applied area do little to clarify the issue because they are designed, not to identify the determinants of human learning, but to provide effective therapy and to change behavior on a pragmatic basis. Clearly, as Skinner had indicated as far back as 1938, whether or not one can extrapolate from animals to humans is a question that can only be resolved by systematic experimental investigation of both animal and human operant behavior.

It is therefore surprising, perhaps, that so little research of this kind has been conducted with human subjects. Rather, the assumption has dominated learning research that while there may be differences of degree or complexity, there are in essence no qualitative differences between human and animal behavior. This *continuity assumption* has taken one of two forms. According to one school of thought, well represented among behavior analysts, the principles of conditioning derived from the study of animal learning are quite sufficient in themselves to account also for human behavior. This essentially zoomorphic view has its anthropomorphic counterpart in cognitive theories that attempt to explain animal performance with concepts derived from observation of human behavior (for a critique of this approach, see Blackman, 1983). Dickinson (1985) for example, argues for a teleological model of both animal and human behavior: "some activities [of animals] are purposive actions controlled by the current value of their goals through knowledge about the instrumental relations between actions and their consequences" (p. 67) and "the knowledge about the action-goal relationship must be encoded in propositional-like form so that it can be acted on by a practical inference process to generate the instrumental performance. In this sense actions are truly rational in a way that responses can never be" (p. 78). Thus it is the case, apparently, that animals have not only cognitive representation but make inferences and formulate propositions (see also Dickinson, 1980; Honig, 1978; Mackintosh & Dickinson, 1979; Roitblat, Bever, & Terrace, 1984).

Although the continuity position in both its forms has held great sway among researchers in the learning area, evidence has been accumulating in recent years that suggests that it is now no longer tenable. Much of this evidence comes from studies of operant behavior in children and adults and it supports the notion that the development of language in humans with the consequent emergence of rule-governed, as

opposed to contingency-shaped, behavior (Skinner, 1966), has a profound effect on behavioral relationships generally. In the light of these findings an alternative theoretical perspective, going beyond either simple continuity or discontinuity assumptions, is now required. It must acknowledge that there is a biological, and indeed psychological, continuity between animals and humans but also that there are qualitative differences (cf. Luria, 1961, 1982; Vygotsky, 1962, 1978). Although the literature on this research has been reviewed in detail elsewhere (Lowe, 1979, 1983), some of the main findings will be briefly summarized here.

A number of studies have shown that the behavior of adult humans on basic schedules of reinforcement differs grossly from animal performance under similar conditions. For example, on the fixed-interval (FI) schedule, where the first response is reinforced after a stated interval has elapsed since the previous reinforcement, adult human behavior bears little resemblance to that of animals and often takes one of two forms—either a continuous and high rate of responding (the *high-rate* pattern) or a very low rate consisting of just one or two responses at the end of the inter-reinforcement interval (the *low-rate* pattern). The FI scallop, and the sensitivity of performance to variations in the schedule value, characteristic of animal behavior, are virtually never seen (Leander, Lippman, & Meyer, 1968; Lowe, 1979; Weiner, 1969; but see also Lowe, Harzem, & Bagshaw, 1978, and Lowe, Harzem, & Hughes, 1978). In addition to these marked differences in performance on the basic schedules, humans also differ from animals in the way they are affected by their previous history of reinforcement. Human subjects exposed to different schedules frequently show a rigidity of performance in the face of altered reinforcement contingencies that is often maladaptive in terms of reinforcement gain or response output, and this too is uncharacteristic of animal behavior.

A series of studies by one of the present authors and colleagues (Bentall, Lowe, & Beasty, 1983, 1985; Lowe, 1983; Lowe, Beasty, & Bentall, 1983) has shown that these distinctive features of adult human operant behavior are absent in preverbal infants, who perform on FI, fixed-ratio (FR) and differential-reinforcement-of-low-rate (DRL) schedules in a manner indistinguishable from that of animals. Taking as an example once more performance on the FI schedule, it was found that human infants show scalloped patterns of responding and sensitivity to the schedule parameters just like that of animals. In contrast, children aged 5 years or older, who have the verbal skills to describe the schedule contingencies to themselves and to formulate rules for responding, show high- or low-rate patterns like adult humans, with similar insensitivity to alterations in schedule value. Children in an intermediate age range of 2½ to 4 years, with less well developed verbal skills, produce neither adult-like nor animal-like patterns of responding, but a highly variable pattern that contains elements of both forms of responding (Lowe *et al.*, 1983; Bentall *et al.*, 1985). Experiments showing that the developmental sequence can be accelerated by appropriate verbal instruction provide further evidence that the acquisition of linguistic skills is the variable responsible for these age-related changes in operant behavior (Lowe, 1983; —and see also Bem, 1967; Luria, 1961).

Recent studies of human choice have revealed major divergences from performance typical of animals. On single and concurrent variable interval (VI) schedules of reinforcement, human behavior, unlike that of animals, does not reliably conform either to Herrnstein's quantitative law of effect or to the generalized matching

law; moreover, the results show that covertly formulated rules for responding are major determinants of human behavior in these settings (Lowe & Horne, 1985; Poppen, 1982; Takahashi & Iwamoto, 1986). The work of Sidman and others on discrimination learning has shown that when conditional discriminations are taught to children of 5 years and older new relations often emerge, for example, symmetry, transitivity, and equivalence (see Sidman & Tailby, 1982). Because these new relations have never been reinforced, they are not encompassed by the three-term contingency (Catania, 1984) and significantly there has been no satisfactory demonstration of these relations in studies of animal learning (Sidman *et al.*, 1982). A series of experiments recently conducted in this laboratory showed that very young children, aged 2 to 3 years old, also failed to form equivalence relations; when taught to label the stimuli, however, most of these children went on to pass the tests for the emergence of equivalence (Lowe, 1986).*

These and other studies, including the work of Vygotsky (1962, 1978), Luria (1961, 1982) and Sokolov (1972), support the following account of human/animal differences in operant conditioning experiments. Without the human capacity for language, animals do not, *pace* Dickinson, form propositions, describe contingencies of reinforcement, or form rules for responding; thus their behavior is affected in very different ways by reinforcement contingencies. On the other hand, models of conditioning that do not take into account the controlling role of verbal behavior, both overt and covert, in human psychology must prove inadequate, whether in the clinical arena or elsewhere.

BEHAVIORAL THEORY IN PRACTICE

Can it really be the case that psychologists who employ behavior modification procedures choose in practice to ignore control by verbal behavior? Presumably, from the moment the client enters the clinician's office and is asked to "sit down" we can begin to collect evidence to the contrary. Indeed, a detailed account of any of the major behavioral interventions commonly employed with linguistically able populations would reveal innumerable practices that appear to be at odds with the conventional animal model of operant conditioning. Two widely known behavior modification procedures, token economies and contingency management, serve to illustrate the point.

TOKEN ECONOMIES

Since the early work of Ayllon & Azrin (1968) with chronic psychiatric patients, token economy programs have been conducted with a variety of clinical and non-clinical subject populations in a variety of different settings (cf. Kazdin, 1977, 1982). This work is frequently cited as evidence for the effectiveness of operant contingencies with humans.

*This study was conducted in collaboration with Allan Beasty.

A standard account of token economies would be that they (a) specify a series of target behaviors for the particular client group, (b) present tokens contingent on the subjects' performance of the target behavior, and (c) allow subjects access to items from a variety of back-up reinforcers through the exchange of tokens (cf. Ayllon & Azrin, 1968). Such a description might apply equally well to a study of animal operant behavior; there is no reference to the role of verbal behavior, either overt or covert. But does the description, in fact, accurately characterize what happens in token economies?

For a number of years, one of the present authors (Higson) conducted a token economy programme with long-stay patients in a psychiatric hospital (Higson, Woods, Tannahill, & Ellis, 1985; Woods, Higson, & Tannahill, 1984). Detailed analysis showed that verbal control was an integral part of this program, for example, staff provided (a) verbal prompts to initiate target behavior, (b) verbal statements, accompanying token presentation, of whether or not the subject's performance of the target behavior matched the specified criteria, (c) verbal descriptions of the contingencies in operation and instructions given to subjects at group meetings, especially when a subject was new to the program, (d) brief written verbal descriptions of the contingencies posted throughout the ward (e.g., "make your bed and earn six tokens"), and (e) a full written description of the contingencies upon each new subject's introduction to the ward. This is entirely characteristic of most token economies with psychiatric patients; extensive use is made of verbal behavior in both initiating and maintaining behavior.

Now it might be argued that this is not the way for a good behaviorist to conduct a token economy program, because it results in the reinforcing contingencies being contaminated by verbal complexities (cf. Michael, 1980) and that instead, one should minimize instructions and concentrate on getting the response-reinforcer relationships correct as is customary in animal experimentation. The evidence suggests that this would be a recipe for failure. In reviews of the token economy literature, Franks & Wilson (1974) and Kazdin (1977) argue that one of the reasons why some clients' behavior is insensitive to the reinforcing contingencies is that therapists' verbal descriptions of the contingencies are not sufficiently detailed or explicit. For example, Franks & Wilson (1974) write:

Instructions combined with reinforcement seem to facilitate performance. . . . The staff concerned have to be well-trained—they must know how best to reinforce behavior, and how to accompany reinforcement with an explicit statement of the contingencies which are operating (e.g., "I gave you four tokens because of the good cleaning job you did this morning").

A number of studies confirm this view. For example, Ayllon & Azrin (1964) found that providing a tangible reinforcer to modify the meal-time behavior of psychiatric patients had no effect on performance unless it was accompanied by instructions that specified the reinforcing contingency; it should also be noted, however, that instructions alone had no enduring effect unless accompanied by reinforcement. Herman & Tramontana (1971) reported that presenting tokens to children as reinforcers for appropriate classroom behavior did not markedly alter behavior until the contingencies were described to the children. Similarly, studies by Suchotliff, Greaves, Stecker, and Berke (1970), Hall, Baker, and Hutchinson (1977) and Baker, Hall,

Hutchinson and Bridge (1977) also testify to the central role of instructions in token economy programs. Interestingly, however, this issue has not received very much attention in subsequent research on token economy programs (Kazdin, 1982).

What it is that controls the behavior of clients in these situations is an obviously critical question: Is behavior under instructional control, or the control of the putative reinforcing contingencies, or some combination of both? As Kazdin (1977) has pointed out, in most programs little attempt has been made to assess the extent to which reinforcement contributes to changes in behavior over and above instructions and yet, in general, little credit has been given to instructions as a factor involved in behavior change. This is exemplified by a report of a token economy program conducted by Nelson and Cone (1979). Token contingencies were introduced to increase the appropriate behavior of psychiatric patients in four different areas: personal hygiene, personal management, ward work, and social skills. Nelson and Cone attribute the observed increase in subjects' performance of the target behaviors entirely to the introduction of the token economy contingencies. In their description of the program, on the other hand, they devote a section to what they term prompts. Here the authors state that

after token reinforcement was initiated for a category of behaviors, verbal instructions, reminders and modelled demonstrations were frequently provided. In addition, posters were placed on the ward walls indicating target behaviors, token values, ward rules, and the ward schedule . . . Observation of 130 instances of subjects' performing the target behaviors during the implementation and probe phases indicated that subjects received some type of individual prompt . . . during 24% of constructive activity, 71% for inappropriate behavior, and 100% for inactivity. (Nelson & Cone, 1979, pp. 260-261)

The plea to get back to instruction-free contingencies in order to achieve pure contingency control (Matthews, Shimoff, Catania, & Sagvolden, 1977; Michael, 1980) might seem appealing. But this aspiration may be founded on a basic misunderstanding of the nature of human behavior, because the evidence suggests that once they have acquired language, humans thenceforth very often respond to contingencies of reinforcement in a way quite different from that of animals. They respond verbally to their responding; they comment on the contingencies to themselves; they reflect; they consider possibilities; they imagine alternatives; they formulate rules. This ongoing commentary on their own behavior and its likely environmental consequences persists regardless of how much the therapist's instructions have been minimized. Indeed, the fewer instructions provided by the therapist the greater is the scope for the influence of the client's own self-instructions, and these may provide a completely erroneous account of the contingencies. It is the influence of such spontaneous, but misleading, formulations of the contingencies and rules for responding that all too often leads to the contingency insensitivity reported in token economy programs (Franks & Wilson, 1974; Kazdin, 1977) and in the experimental literature on human operant behavior (cf. Lowe, 1979, 1983; Lowe & Horne, 1985).

CONTINGENCY MANAGEMENT

The effects of verbal behavior can also be observed in various *contingency management* procedures which involve the systematic scheduling of both positive and negative consequences for behavior. One form of contingency management is achieved

through the use of *contingency contracting*. This procedure involves the negotiation of a written contract between two or more individuals clearly specifying the target behaviors agreed on by each person, with specially arranged consequences attendant upon successful performance of these behaviors. The contract may also specify setting conditions for the target behavior and the consequences of noncompliance with the terms of the contract. Contingency contracting has been successfully employed in a variety of clinical settings and with a variety of problems, for example, marital problems (Crowe, 1978; Stuart, 1969), child delinquency (Stuart, 1971; Tharp & Wetzel, 1969), classroom management (Homme, Csanzi, Gonzales, & Rechs, 1970) alcohol abuse (Miller, 1972), and obesity (Foreyt, 1977; Mann, 1972, 1977). The following two examples may serve to illustrate some of the factors at work in standard contingency management and contracting procedures.

The contingency-management approach adopted with problem drinkers typically provides either positive consequences for a reduction in the rate of drinking or the amount of alcohol drunk, or negative consequences for the occurrence of alcohol drinking, or some combination of both. Cohen, Liebson, and Faillace (1971), for example, describe a series of studies conducted with a 39-year-old, hospitalized chronic alcoholic with a 10-year history of alcohol abuse. The target behavior was a reduction in the overall amount of alcohol drunk each day, for which a positive or negative consequence was presented according to whether the subject drank more or less than the specified limit. In one study a free-operant drinking phase was instituted in which the subject had access to 24 ounces of 95 proof ethanol each day. During a contingent reinforcement phase, if the subject drank 5 ounces of alcohol or less on a particular day, he was placed in an enriched ward environment that provided the opportunity to work for money and for access to private telephone, recreation room, and television. If, on the other hand, the subject drank over 5 ounces he was placed in an impoverished environment (loss of all privileges) for the rest of the day. During control conditions no contingencies were in operation. The results of this single-case study indicated that controlled drinking (under 5 ounces per day) was maintained for as long as 5 weeks during contingent phases, with a return to excessive drinking during noncontingent phases. Similar findings were reported in other studies that have used the same type of procedure with groups of alcohol abusers (Cohen, Liebson, & Faillace, 1972, 1973).

A study by Mann (1977) involving weight-reduction in male and female subjects provides an example of contingency contracting. The contract (a) required each subject to surrender a large number of items considered to be of personal value, (b) required the subject to be weighed regularly, (c) prescribed the manner in which the subject could earn back or permanently lose his or her valuables (i.e., a statement of the contingencies), and, (d) stipulated that the researcher, at his discretion, would change the procedures from baseline, to treatment, back to baseline, and back to treatment conditions (a single-subject design was used). Three forms of reinforcement contingency were specified in the contract: (a) as soon as each 2 lb weight reduction was achieved the subject received one valuable; (b) the subject was presented with a bonus valuable for losing a minimum number of pounds by the end of each successive 2-week period during the treatment condition; (c) some of the valuables were delivered to the subject only if and when the target weight requirement (specified at the outset) was met. In addition, if the subject decided at any time to opt out of

the program the researcher kept possession of all the remaining valuables. Mann reported that the contract procedure was successful in producing significant reductions in weight for all subjects.

Although these examples together with many other studies show that contingency management procedures can be successful, they also cast doubt on the role of programmed reinforcement. For, as Michael (1980) has previously observed, in cases such as these the behavior being affected is so distanced in time from its programmed consequences that it cannot possibly be directly reinforced by them. In animal operant research even very short delays between the operant response and the presentation of the reinforcer (i.e., ranging from a few seconds to a few minutes) can seriously retard or eliminate the acquisition of behavior (Davey, 1981; Skinner, 1938). In many contingency management studies, however, including the two examples presented earlier, the delay between the occurrence of the behavior to be reinforced and the putative reinforcer may be several hours, or even days, long. We would agree with Michael (1980) when he suggests that "Such effects are probably always mediated through some form of rule statement or rule control, which is typically not mentioned or analysed." These rule statements will incorporate the instructions given to the subject or, as in the case of contingency contracting, the written descriptions of the contingencies. But where our account differs from that of Michael is that he cites such cases as being exceptional; it is assumed that when behavior is followed closely in time by a particular consequence that it will be free of rule statements and rule control. There can, however, be no good grounds, theoretical or empirical, for this assumption. On the contrary, apart from those situations where they act "without awareness" (Hefferline, Keenan, & Harford, 1959; Laurenti-Lions, Gallego, Chambille, Vardon, & Jacquemin, 1985), it seems that humans, having once acquired language, will persist in using it to construe their environment whatever may be the temporal relationship between responding and its consequences.

CONCLUSION

Close analysis of most behavior modification studies yields the startling conclusion that, in spite of what is commonly supposed, changes in target behavior cannot be solely or directly the result of reinforcing contingencies; in order for the programmed contingencies to be effective the behavior must, to some extent, be appropriately rule governed. Yet in these studies, the verbal behavior of therapist and client alike is all too often neither mentioned nor analyzed. If rule-governed behavior is as pervasive and potent a variable in human learning as the evidence from basic research and applied work now suggests then it must surely be folly for applied behavior analysis to ignore it (and see Bentall, Lowe, & Higson, *in press*; Woods & Lowe, *in press*; Zettle & Hayes, 1982).

This, of course, raises the question of whether Skinner's radical behaviorist theory is in fact equal to coping with the complexities of human behavior, including the effects of verbal behavior, much of which occurs covertly. Theoretical objections to the study of covert events, largely on the grounds that there can be no public

agreement about their validity, has come from methodological behaviorism, a theoretical position which has been adopted by many behavior therapists. But Skinner has consistently argued against this view, claiming that it misguidedly adheres to the outmoded tenets of logical positivism and operationism. Indeed the principal distinguishing feature of his radical behaviorism is that it considers that a science of behavior, like other sciences, must deal with events that are not directly observable; inference, therefore, is held to be essential in the study of behavior, regardless of parsimony (Skinner, 1945, 1953, 1963, 1974, 1984).

Thus it is surely a strange irony of contemporary psychology that an approach that, as far back as 1945, established its identity on the basis of its recognition of the inner life of humans should so often be charged with the error of ruling it out of court. It is widely asserted, for example, that Skinner's is a "black box" account of human behavior, that it does not deal with consciousness and cognitive processes, that it eschews the analysis and modification of private events, and that it shuns inferential accounts of behavior because they are unparsimonious (see Chomsky, 1975; Harré & Secord, 1972; Kendall & Hollon, 1979; Koestler, 1967; Ledgidge, 1978; Locke, 1979; Mahoney, 1977; Wilson, 1978). For instance, within clinical psychology the exponents of cognitive behavior therapy have found it necessary to adopt the conceptual apparatus of cognitivism apparently out of a belief that the behavioral approach cannot deal with the modification of peoples' covert behavior (cf. Lowe & Higson, 1981, 1983). It is at least partly the responsibility of behaviorists themselves that such misconceptions about radical behaviorism are so widespread. For unhappily, despite the theoretical lead given by Skinner, radical behaviorists until recently have been reluctant, in both basic research and applied work, to investigate the role of language in human learning. Although Skinner's account of the role of verbal behavior in the development of human consciousness is similar in many respects to that of Vygotsky (1962) and Luria (1961), it has not had anything like a comparable impact on psychological research. Instead the operant approach has, until recently, been concerned almost exclusively with animal behavior or with human behavior treated as if it did not differ significantly, in terms of controlling variables, from the key-peck of the pigeon or the lever-press of the rat. There is now, however, a rapidly growing research interest in the factors that distinguish human from animal operant behavior and in the role of verbal behavior in bringing about these differences (e.g., Catania, Matthews, & Shimoff, 1982; Hayes, Brownstein, Zettle, Rosenfarb, & Korn, 1986; Matthews, Catania, & Shimoff, 1985; Vaughan, 1985; Sidman *et al.*, 1982; Wearden & Shimp, 1985). Until a similar consideration is given to such issues in applied behavior analysis the potential contribution of operant theory to clinical psychology will to a significant degree remain unrealized.

REFERENCES

- Ayllon, T., & Azrin, N. H. (1964). Reinforcement and instructions with mental patients. *Journal of the Experimental Analysis of Behavior*, 7, 327-331.
- Ayllon, T., & Azrin, N. H. (1968). *The token economy: A motivational system for therapy and rehabilitation*. New York: Appleton-Century-Crofts.

- Baker, R., Hall, J. N., Hutchinson, K., & Bridge, G. (1977). Symptom changes in chronic schizophrenic patients on a token economy: A controlled experiment. *British Journal of Psychiatry*, *131*, 381–393.
- Beech, H. R. (1974). Behaviour therapy. In P. Morris (Ed.), *Behaviourism*. (pp. 85–115) Milton Keynes, UK: The Open University Press.
- Bem, S. L. (1967). Verbal self-control: The establishment of effective self-instruction. *Journal of Experimental Psychology*, *74*, 485–491.
- Bentall, R. P., Lowe, C. F., & Beasty, A. (1983). Does language produce behavioural rigidity in humans? *Behaviour Analysis Letters*, *3*, 251.
- Bentall, R. P., Lowe, C. F., & Beasty, A. (1985). The role of verbal behavior in human learning II: Developmental differences. *Journal of the Experimental Analysis of Behavior*, *43*, 165–181.
- Bentall, R. P., Lowe, C. F., & Higson, P. (1984). Teaching verbal self-regulation to chronic schizophrenic patients: Efficacy and generalisation. *Behavioural Psychotherapy*, *15*, 58–76.
- Blackman, D. E. (1983). On cognitive theories of animal learning: Extrapolation from humans to animals? In G. C. L. Davey (Ed.), *Animal models of human behavior* (pp. 37–50). Chichester, UK: Wiley.
- Bloomfield, T. M. (1976). About Skinner: Notes on the theory and practice of “Radical Behaviourism.” *Philosophy of the Social Sciences*, *6*, 75–82.
- Boulougouris, J. C. (Ed.). (1982). *Learning theory approaches to psychiatry*. Chichester, UK: Wiley.
- Branch, M. N., & Malagodi, E. F. (1980). Where have all the behaviorists gone? *The Behavior Analyst*, *3*, 31–38.
- Brewer, W. P. (1974). There is no convincing evidence for operant or classical conditioning in adult humans. In W. B. Weiner & D. S. Palermo (Eds.), *Cognition and the symbolic processes* (pp. 1–42). Hillsdale, NJ: Erlbaum.
- Catania, A. C. (1984). *Learning* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Catania, A. C., Matthews, B. A., & Shimoff, E. (1982). Instructed versus shaped human verbal behavior: Interactions with nonverbal responding. *Journal of the Experimental Analysis of Behavior*, *38*, 233–248.
- Chomsky, N. (1975) *Reflections on language*. Glasgow, UK: William Collins.
- Cohen, M., Liebson, I. A., & Faillace, L. A. (1971). The role of reinforcement contingencies in chronic alcoholism: An experimental analysis of one case. *Behavior Research and Therapy*, *9*, 375–379.
- Cohen, M., Liebson, I. A., & Faillace, L. A. (1972). A technique for establishing controlled drinking in chronic alcoholics. *Diseases of the Nervous System*, *33*, 46–49.
- Cohen, M., Liebson, I. A., & Faillace, L. A. (1973). Controlled drinking by chronic alcoholics over extended periods of free access. *Psychological Reports*, *32*, 1107–1110.
- Crowe, M. J. (1978). Behavioural approaches to marital and family problems. In R. N. Gaince & B. L. Hudson (Eds.), *Current themes in psychiatry 1*. (pp. 257–274). London: Macmillan.
- Davey, G. C. L. (1981). *Animal learning and conditioning*. London: Macmillan.
- Dickinson, A. (1980). *Contemporary animal learning theory*. Cambridge, UK: Cambridge University Press.
- Dickinson, A. (1985). Actions and habits: The development of behavioral autonomy. *Philosophical Transactions of the Royal Society, London*, 67–68.
- Dietz, S. M. (1978). Current status of applied behavior analysis: Science versus technology. *American Psychologist*, *33*, 805–814.
- Epling, W. F., & Pierce, W. D. (1986). The basic importance of applied behavior analysis. *The Behavior Analyst*, *9*, 89–100.
- Foreyt, J. P. (1977). *Behavioral treatments of obesity*. New York: Pergamon Press.
- Franks, C. M., & Wilson, G. T. (1974). *Annual review of behavior therapy and practice*. New York: Brunner/Mazel.
- Hall, J. N., Baker, R. D., & Hutchinson, K. (1977). A controlled evaluation of token economy procedures with chronic schizophrenic patients. *Behavior Research and Therapy*, *15*, 261–283.
- Harré, R., & Secord, P. F. (1972). *The explanation of social behaviour*. Oxford: Blackwell.
- Hayes, S. C., Rincover, A., & Solnick, J. (1980). The technical drift of applied behavior analysis. *Journal of Applied Behavior Analysis*, *13*, 275–286.
- Hayes, S. C., Brownstein, A. J., Zettle, R. D., Rosenfarb, I., & Korn, Z. (1986). Rule-governed behavior and sensitivity to changing consequences of responding. *Journal of the Experimental Analysis of Behavior*, *45*, 237–256.
- Hefferline, R. F., Keenan, B., & Harford, R. A. (1959). Escape and avoidance conditioning in human subjects without their observation of the response. *Science*, *130*, 1338–1339.

- Herman, S., & Tramontana, J. (1971). Instructions and group versus individual reinforcement in modifying disruptive group behaviour. *Journal of Applied Behavior Analysis*, 4, 113–119.
- Higson, P. J., Woods, P. A., Tannahill, M. M., & Ellis, N. C. (1985). The role of meals as a reinforcing event in a token economy programme. *British Journal of Psychiatry*, 147, 170–174.
- Homme, L., Csanzi, A., Gonzales, M., & Rechs, J. (1970). *How to use contingency contracting in the classroom*. Champaign, IL: Research Press.
- Honig, W. K. (1978). Studies of working memory in the pigeon. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior* (pp. 1–14). Hillsdale, NJ: Erlbaum.
- Kazdin, A. E. (1977). *The token economy: A review and evaluation*. New York: Plenum Press.
- Kazdin, A. E. (1982). The token economy: A decade later. *Journal of Applied Behavior Analysis*, 15, 431–445.
- Kendall, P. C., & Hollon, S. D. (1979). *Cognitive-behavioral interventions: Theory, research and procedures*. New York: Academic Press.
- Koestler, A. (1967) *The ghost in the machine*. London: Hutchinson.
- Kohler, F. W., & Greenwood, C. R. (1986). Towards a technology of generalisation: The identification of natural contingencies of reinforcement. *The Behavior Analyst*, 9, 19–26.
- Laurenti-Lions, L., Gallego, J., Chambille, B., Vardon, G., & Jacquemin, C. (1985). Control of myoelectrical responses through reinforcement. *Journal of the Experimental Analysis of Behavior*, 44, 185–193.
- Leander, J. D., Lippman, L. G., & Meyer, M. E. (1968). Fixed-interval performance as related to subjects' verbalizations of the reinforcement contingency. *Psychological Record*, 18, 469–474.
- Ledgidge, B. (1978). Cognitive-behavior modification: A step in the wrong direction? *Psychological Bulletin*, 85, 353–375.
- Locke, E. A. (1979). Behaviour modification is not cognitive and other myths: A reply to Ledgidge. *Cognitive Therapy and Research*, 3, 119–125.
- Lowe, C. F. (1979). Determinants of human operant behaviour. In M. D. Zeiler & P. Harzem (Eds.), *Advances in analysis of behaviour (Vol. 1): Reinforcement and the organisation of behaviour* (pp. 159–192). Chichester, UK: Wiley.
- Lowe, C. F. (1983). Radical behaviourism and human psychology. In G. C. L. Davey (Ed.), *Animal models and human behaviour: Conceptual, evolutionary and neurobiological perspectives* (pp. 71–93). Chichester, UK: Wiley.
- Lowe, C. F. (1986). *The role of verbal behavior in the emergence of equivalence classes*. Paper presented to the Twelfth Annual Convention of the Association for Behavior Analysis, Milwaukee, WI.
- Lowe, C. F., Beasty, A., & Bentall, R. P. (1983). The role of verbal behavior in human learning: Infant performance on fixed-interval schedules. *Journal of the Experimental Analysis of Behavior*, 39, 157–164.
- Lowe, C. F., Harzem, P., & Bagshaw, M. (1978). Species differences in temporal control of behavior II: Human performance. *Journal of the Experimental Analysis of Behavior*, 29, 351–361.
- Lowe, C. F., Harzem, P., & Hughes, S. (1978). Determinants of operant behaviour in humans: Some differences from animals. *Quarterly Journal of Experimental Psychology*, 30, 373–386.
- Lowe, C. F., & Higson, P. J. (1981). Self-instructional training and cognitive behaviour modification: A behavioural analysis. In G. C. L. Davey (Ed.), *Applications of conditioning theory* (pp. 162–188). London: Methuen.
- Lowe, C. F., & Higson, P. J. (1983). Is all behaviour modification “cognitive”? In E. Karas (Ed.), *Current issues in clinical psychology* (pp. 207–227). London: Plenum Press.
- Lowe, C. F., & Horne, P. J. (1985). On the generality of behavioural principles: Human choice and the matching law. In C. F. Lowe, M. Richelle, D. E. Blackman, & C. M. Bradshaw (Eds.), *Behaviour analysis and contemporary psychology* (pp. 97–115). London: Erlbaum.
- Luria, A. R. (1961). *The role of speech in the regulation of normal and abnormal behavior*. New York: Liveright.
- Luria, A. R. (1982). *Language and cognition*. New York: Wiley.
- Mackintosh, N. J., & Dickinson, A. (1979). Instrumental (type 11) conditioning. In A. L. Dickinson and R. A. Boakes (Eds.), *Mechanisms of learning and motivation: A memorial volume to Jerzy Konorski* (pp. 143–169). Hillsdale, NJ: Lawrence Erlbaum Associates, 1979.
- Mahoney, M. J. (1977). Reflections on the cognitive-learning trend in psychotherapy. *American Psychologist*, 32, 5–13.
- Mann, R. A. (1972). The behaviour-therapeutic use of contingency contracting to control an adult behaviour problem: Weight control. *Journal of Applied Behavioral Analysis*, 5, 99–109.

- Mann, R. A. (1977). The behaviour-therapeutic use of contingency contracting to control an adult behaviour problem: Weight control. In J. P. Foreyt (Ed.), *Behavioral treatments of obesity* (pp. 122–143). New York: Pergamon Press.
- Matthews, B. A., Catania, A. C., & Shimoff, E. (1985). Effects of uninstructed verbal behavior on nonverbal responding: Contingency descriptions versus performance descriptions. *Journal of the Experimental Analysis of Behavior*, *43*, 155–164.
- Matthews, B. A., Shimoff, E., Catania, A. C., & Sagvolden, T. (1977). Uninstructed human responding: Sensitivity to ratio and interval contingencies. *Journal of the Experimental Analysis of Behavior*, *27*, 453–467.
- Michael, J. L. (1980). Flight from behavior analysis. *The Behavior Analyst*, *3*, 1–21.
- Michael, J. L. (1985). Fundamental research and behaviour modification. In C. F. Lowe, M. Richelle, D. E. Blackman, & C. M. Bradshaw (Eds.), *Behaviour analysis and contemporary psychology* (pp. 159–164). London: Erlbaum.
- Miller, P. M. (1972). The use of behavioural contracting in the treatment of alcoholism: A case report. *Behaviour Therapy*, *3*, 593–596.
- Nelson, G. L., & Cone, J. D. (1979). Multiple-baseline analysis of a token economy for psychiatric patients. *Journal of Applied Behavior Analysis*, *12*, 255–271.
- Poppen, R. (1982). Human fixed-interval performance with concurrently programmed schedules: A parametric analysis. *Journal of the Experimental Analysis of Behavior*, *37*, 251–266.
- Roitblat, H. L., Bever, T. G., & Terrace, H. S. (Eds.). (1984). *Animal cognition*. Hillsdale, NJ: Erlbaum.
- Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, *37*, 23–44.
- Sidman, M., Raizin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discrimination of rhesus monkeys, baboons, and children. *Journal of the Experimental Analysis of Behavior*, *37*, 5–22.
- Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.
- Skinner, B. F. (1945). The operational analysis of psychological terms. *Psychological Review*, *52*, 270–277.
- Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review*, *57*, 193–216.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
- Skinner, B. F. (1963). Behaviorism at fifty. *Science*, *134*, 566–602.
- Skinner, B. F. (1966). An operant analysis of problem solving. In B. Kleinmuntz (Ed.), *Problem solving: Research, method and theory* (pp. 225–257). New York: Wiley.
- Skinner, B. F. (1974). *About behaviorism*. London: Jonathan Cape.
- Skinner, B. F. (1984). Coming to terms with private events. In A. C. Catania & S. Harnad (Eds.), *The Behavioral and Brain Sciences*, *7*, 572–581.
- Sokolov, A. N. (1972). *Inner speech and thought*. New York: Plenum Press.
- Stuart, R. B. (1969). Operant interpersonal treatment for marital discord. *Journal of Consulting and Clinical Psychology*, *33*, 675–682.
- Stuart, R. B. (1971). Behavioural contracting within the families of delinquents. *Journal of Behaviour Therapy and Experimental Psychiatry*, *2*, 1–11.
- Suchotliff, L., Greaves, S., Stecker, H., & Berke, R. (1970). Critical variables in the token economy system. *Proceedings of the 78th Annual Convention of the American Psychological Association*, *5*, 517–518.
- Takahashi, M., & Iwamoto, T. (1986). Human concurrent performances: The effects of experience, instructions, and schedule-correlated stimuli. *Journal of the Experimental Analysis of Behavior*, *45*, 257–267.
- Tharp, R. G., & Wetzel, R. T. (1969). *Behaviour modification in the natural environment*. New York: Academic Press.
- Vaughan, M. E. (1985). Repeated acquisition in the analysis of rule-governed behavior. *Journal of the Experimental Analysis of Behavior*, *44*, 175–184.
- Vygotsky, L. S. (1962). *Thought and language*. New York: Wiley.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.). Cambridge, MA: Harvard University Press.
- Wearden, J. H., & Shimp, C. P. (1958). Local temporal patterning of operant behavior in humans. *Journal of the Experimental Analysis of Behavior*, *44*, 315–324.
- Weiner, H. (1969). Controlling human fixed-interval performance. *Journal of the Experimental Analysis of Behavior*, *12*, 349–373.

- Wilson, T. (1978). Cognitive behavior therapy: Paradigm shift or passing phase? In J. Foreyt & D. Rathjen (Eds.), *Cognitive behavior therapy: Research and application*. New York: Plenum Press.
- Woods, P. A., Higson, P. J., & Tannahill, M. M. (1984). Token economy programmes with chronic psychotic patients: The importance of direct measurement and objective evaluation for long-term maintenance. *Behaviour Research and Therapy*, 22, 41-51.
- Woods, P. A., & Lowe, C. F. (in press). Verbal control of inappropriate social behaviour in mentally handicapped adults. In J. Berg (Ed.), *Science and service in mental retardation*. London: Methuen.
- Zettle, R. D., & Hayes, S. C. (1982). Rule-governed behavior: A potential theoretical framework for cognitive therapy. In P. C. Kendall (Ed.), *Advances in cognitive-behavioral research and therapy* (Vol. 1, pp. 76-118). New York: Academic Press.

CHAPTER 8

Language Conditioning Clinical Issues and Applications in Behavior Therapy

Georg H. Eifert

INTRODUCTION

Language is undoubtedly one of the most powerful means by which human behavior is controlled, and most schools of psychotherapy—including behavior therapy—rely on some form of language-based interventions as part of their therapeutic methods. Although animals possess rudimentary forms of language (e.g., the chimp's sign language), it is generally agreed that complex language is one of the key characteristics that distinguishes humans from animals. Razran (1965) pointed out that a system of psychology based wholly on conditioned reflexes, not drawing a basic distinction between animal and human learning, would be highly mechanistic and reductionistic. Pavlov, in his later years, clearly recognized this danger. He viewed speech as a system of second signals—unique to human beings—that are in essence abstractions of reality and means of generalization. Although he regarded words as conditioned stimuli governed by “the fundamental laws of learning,” he also emphasized the differences between words and other stimuli: “Of course a word is for a man as much a real conditioned stimulus as are other stimuli common to men and animals, yet at the same time it is so all-comprehending that it allows no quantitative or qualitative comparisons with conditioned stimuli in animals” (in Razran, 1965, p. 48).

Pavlov's distinction is indeed crucial: whereas there may be much overlap between animal and human learning at the level of basic learning principles, there are also important qualitative and not just quantitative differences between animal and human learning. These differences are largely the result of human ability to use language and symbols. Furthermore, in order to account for the complexities and flexibility of human behavior, it is also necessary to add a personality-theory level

to basic learning principles. Although influential behaviorists (e.g., Eysenck, 1960; Staats, 1963) have long recognized this necessity, behavior therapists have somewhat neglected interindividual differences in their theory-building efforts (Ross, 1985).

Literally hundreds of studies have been conducted investigating the effectiveness of language conditioning and the variables and parameters it involves. A computer literature search in preparation for this chapter yielded 525 laboratory and clinical studies carried out between 1967 and 1984. Reviews of studies conducted prior to 1967 can be found in Kanfer (1968) and Krasner (1958), as well as in an excellent book by Paivio (1971) entitled *Imagery and Verbal Processes*. I noted in my literature search that interest in studying language conditioning has sharply declined in recent years. For instance, most of the studies (71%) were conducted in the first half of the review period with progressively fewer studies reported in recent years.

The purpose of this chapter is not to bore the reader with an exhaustive (and exhausting!) review of all these studies, but to examine the results of those studies investigating the emotion-eliciting, reinforcing, and behavior-directive functions of language, because they are of particular relevance for behavior therapy. It will become clear that it is not the topic, type, or results of studies *per se* that is to blame for the growing lack of interest in this area. One of the factors responsible for this development is a change in zeitgeist favoring more cognitive theories (see Mahoney, 1977); another is the lack of a unifying conceptual framework that would have been necessary to grasp the full theoretical and clinical significance of the bulk of findings; and finally, there has been a lack of concern for integrating and linking those findings with other developments in behavior therapy.

TERMINOLOGY AND BASIC CONCEPTS

Let me first briefly define the various terms and techniques used in experiments and clinical studies investigating language conditioning. The term *verbal conditioning* generally refers to the operant conditioning of verbal behavior, that is, the modification of verbal behavior by means of generalized conditioned verbal and nonverbal reinforcers, such as approval or disapproval (cf. Krasner, 1958). Typically, therapists have attempted to increase or decrease the frequency of certain types of client verbal responses (e.g., self-referent affective statements, the use of "I" or "we") by expressing approval verbally (e.g., "right," "good"), or nonverbally by head-nodding, smiling, or leaning forward. More sophisticated methods of reinforcement have combined approval with empathic reflection in a client-centered fashion.

The term *semantic conditioning* was coined by Razran (1939), who described it as a classical conditioning to meaning, that is, the conditioning of a response, for example, galvanic skin response (GSR) to the meaning of a word or sentence. In a typical semantic conditioning experiment an unconditioned stimulus (UCS) (e.g., food) might be paired with the sight of the word *hare*. After several pairings subjects would salivate to the sight of the word *hare* without UCS reinforcement. In subsequent generalization trials with adults one would expect to find a semantic rather than a phonetic generalization of the conditioned response (CR). In other words, the CR is likely to generalize along a meaning dimension of semantically related words such

as *rabbit* or *bunny* rather than to words that sound similar or even the same (e.g., hair). In a series of studies with children, Drinkwater (1968, 1972) found that generalization is stronger to high- than to low-affect verbal stimuli, and that older children of normal intelligence typically generalize along a semantic dimension, whereas the responses of mentally handicapped children generalize more to phonetically similar words. Maltzman (1977) has repeatedly used innocuous stimuli, such as tones or lights, as UCS to condition GSR changes to words. Furthermore, numerous studies on higher-order conditioning (see Razran, 1971; Staats, 1968, 1975) have demonstrated that emotive words can function as UCS and transfer their meaning onto other stimuli, as well as elicit changes in GSR responding and overt motor behavior.

As will be described in the next section, the processes involved in semantic classical conditioning and verbal operant conditioning are inextricably intertwined and the distinction between them is therefore somewhat arbitrary. This is the main reason why recently some researchers have preferred to use the more general term *language conditioning* to indicate the interrelatedness of the emotion-eliciting, reinforcing, and behavior-directive functions of language.

LANGUAGE CONDITIONING IN THE SOCIAL-BEHAVIORIST PARADIGM

LANGUAGE CONDITIONING AND THE FUNCTIONS OF LANGUAGE

According to the social-behaviorist paradigm (Staats, 1968, 1975), salient environmental stimuli have three functions. (a) They elicit affective-emotional responses in the individual on an unlearned or learned basis (A-function). These are central nervous system responses with peripheral physiological indexes, such as smooth muscle responding in the viscera, glandular responses, and so on. (b) Consequently, affective stimuli will act as reinforcers in both operant and classical conditioning situations (R-function). (c) In addition, affective stimuli will also direct overt behavior through learning (D-function): we learn to approach stimuli that elicit positive emotional responses and to avoid stimuli that elicit negative emotional responses. The affective, reinforcing, and directive functions of stimuli are related and depend upon the hedonic value of the stimulus.

One of the significant qualitative differences between animal and human learning that Pavlov referred to is that language serves important symbolic functions by providing humans with emotional experiences without exposure to the actual physical stimuli or events. By means of classical conditioning, humans acquire a verbal-emotional repertoire consisting of a large number of words that come to be emotional stimuli. These verbal stimuli (or labels) and images will acquire the same A-R-D functions as the objects or situations they refer to.

Moreover, a personality level of theory is necessary to account for a person's actual behavior and interindividual differences. This conception indicates that individuals learn complex repertoires of behavior (personality repertoires) over long periods of time. Such learning is cumulative and hierarchical, with the acquisition of one repertoire leading into, and sometimes becoming part of, a more complex

repertoire. The most central of these personality repertoires is the emotional-motivational system with the two other systems, the language-cognitive and the sensorimotor, developing from it. There is much overlap and interaction between the three behavioral repertoires through a number of important subrepertoires (e.g., verbal-emotional, verbal-motor; see also Figure 1).

LANGUAGE CONDITIONING IN ANXIETY AND DEPRESSION

Rather than focusing on research with nonsense syllables and the like, I will attempt to show the clinical relevance of language conditioning by examining its role in the origin and maintenance of such pervasive clinical problems as anxiety disorders and depression. These will be used as convenient examples and frequently referred to in subsequent sections on the role of language conditioning in prominent cognitive-behavioral intervention techniques. The advantage of social-behaviorist theory is that it specifies the psychological principles by which certain thoughts and self-statements come to elicit negative emotional responses and avoidance behavior.

Anxiety. The major implication of social-behaviorist theory for understanding anxiety disorders is that it is not necessary for an individual to have an actual traumatic experience to develop a phobia. The association of inappropriate and/or negative emotion-eliciting verbal stimuli (labels) with certain objects or situations is sufficient for those objects and situations to acquire aversive properties. Hence phobias with no history of overt aversive conditioning could have been acquired vicariously and/or by means of semantic classical conditioning. As the emotion-eliciting, reinforcing, and directive functions of emotive verbal stimuli are interrelated, it means that following conditioning a phobic stimulus will not only elicit a negative emotional response (anxiety) but also lead to avoidance behaviors.

For example, agoraphobic clients constantly pair negative thoughts and verbal stimuli with images of panic and disaster in potentially frightening situations. This means that agoraphobic persons do not need to have direct reconditioning experiences to remain phobic and continue to avoid these situations. They may condition themselves by providing their own verbal-symbolic stimuli, which elicit negative emotional responses. If an agoraphobic woman imagines having a panic attack in the supermarket, with the consequent embarrassment of other people thinking she was ill or "going crazy," these images and thoughts are sufficient to induce anticipatory fear and the avoidance of that particular situation, even though she may never have actually experienced a panic attack in a supermarket.

Depression. Cognitive therapists (e.g., Beck, 1976) have repeatedly emphasized the importance of negative thoughts and self-statements in depression without, however, specifying the psychological principles by which these thoughts and statements influence mood and behavior. In their recent social-behaviorist account of depression, Staats and Heiby (1985) explained the role of verbal and semantic (self-)conditioning in the origin and maintenance of depression. They argue that the depressive's dysphoric state represents an enduring negative emotional response that elicits certain verbal and other motor behavior. For instance, the experience of this negative emotional state tends to elicit a large class of words that have a negative meaning; words that clients often use to describe feelings of worthlessness, thoughts of suicide, and

so on. The dysphoric state also produces difficulty in thinking and concentrating because it elicits those negative meaning words so strongly that they interfere with the trains of language responses that would otherwise occur in response to the ordinary affairs of life. The word groups that the negative emotional state elicits are a function of the individual's learned language-cognitive repertoires and are commonly called opinions, beliefs, and so on. For instance, if a man loses his job he may either blame society or the economic system, or he may blame himself—in which case the result may be a deep and lasting dysphoria with subsequent feelings of inadequacy, self-depreciation, loss of self-esteem, and guilt. In other words, whether or not people will say self-recriminatory things and produce negative emotions in themselves depends on their learned language-cognitive repertoires. In this respect, the social-behaviorist paradigm and cognitive theories of depression are not in contrast: both approaches assume that belief systems mediate how persons interpret the events involved. Once these personality systems have been learned, they may cause emotions and motor behavior—this is why Staats (1975) considers personality both an effect and a cause.

The problem of low self-esteem in depressive persons can be regarded as a lack of positive self-reinforcement. Staats and Heiby (1985) pointed out that positive self-reinforcement depends on a “rich positive verbal-emotional repertoire,” that is, a large number of self-referring statements that elicit positive emotional responses:

Some people . . . richly supply themselves with positive affective-reinforcing-directive (A-R-D) stimuli and others do not. . . . A dysfunctionally low level of [positive] self-administered emotional stimuli may be a consequence of inappropriate, learned, negative self-evaluations, which may be associated with a memory deficit of positive events in the language-cognitive repertoire. (pp. 311–312)

There may also be actual skill dysfunctions in the language-cognitive and sensorimotor repertoires that additionally lead to negative self-labeling, that is, a high level of self-administered negative stimuli. This account highlights the complexities of language conditioning in human functioning and shows that language conditioning is not an esoteric mechanistic laboratory technique but a powerful and versatile learning principle involved in many forms of complex human (dys)functions. The following sections will therefore examine the role of language conditioning in the treatment of such dysfunctions.

APPLICATIONS OF LANGUAGE CONDITIONING IN BEHAVIOR THERAPY

Two of the most important criteria for evaluating a theory are the quality of the experimental and clinical data base supporting it, and its heuristic value in clinical application. A considerable number of general psychological experiments showed that the affective value of positively and negatively toned words can be transferred to either neutral stimuli (e.g., nonsense syllables) or existing attitudinal stimuli (e.g., nationality names). It was demonstrated that language conditioning led to physiological changes as well as changes in overt behavior toward these objects. These experiments and procedures have been described and discussed elsewhere (Staats, 1968, 1975). It was also suggested that the same principles and procedures should

be applicable for the treatment of actual clinical problems. However, studies examining the clinical efficacy of purely language-based interventions—such as semantic, verbal, and covert conditioning interventions—have produced somewhat more mixed results. Unfortunately, limitation of space makes it necessary to exclude covert conditioning techniques (see Cautela, 1973) from this discussion. I should briefly mention though that these techniques have not fulfilled the high expectations aroused by them when they were introduced in the early 1970s. O’Leary and Wilson (1980) pointed out that whereas covert sensitization has been shown to be effective in the treatment of specific forms of sexual deviance, it has proved to be no more effective than placebo factors in the treatment of obesity. Similarly, in a review of behavioral treatments of smoking, Brownell (1984) concluded that the weight of evidence argues against the use of covert sensitization.

In the following sections I will first review the findings from studies in the more promising areas of semantic and verbal conditioning interventions and then discuss cognitive-behavioral interventions from a social-behaviorist language conditioning perspective.

SEMANTIC AND VERBAL CONDITIONING INTERVENTIONS

Hekmat and his co-workers conducted a series of studies investigating the effects of various types of semantic desensitization techniques on different phobias. Initially, these studies closely followed the experimental procedure of Staats’ early studies. Phobic word stimuli (snake or spider) were paired with positive evaluative words (e.g., gift, vacation, joy). In the first study (Hekmat & Vanian, 1971) subjects merely had to listen to the word pairs and imagine the positive words as clearly and vividly as they could. In the second study (Hekmat, 1972) one group again listened to the word pairs and imagined the positive words, whereas the other treatment group had to say aloud the positive evaluative word after the phobic word was presented to them. Both studies included a control group in which a phobia-irrelevant word (peach) was paired with positive evaluative words. The results of both studies were very encouraging: following therapy, clients in the treatment groups evaluated the phobic animals much less negatively on semantic differential ratings and achieved near-perfect approach scores in behavior tests.

Hekmat (1977) modified the semantic desensitization procedure to include positive imagery and image-induced relaxation. Words themselves no longer served as counterconditioning agents, that is, positive evaluative words were not directly paired with the word *snake* but were only used to induce a pleasant image or scene. These scenes were then paired with anxiety-provoking scenes elicited by the therapist saying the word *snake*. The results show a reduction of self-reported snake anxiety, positive changes in affective evaluation, and increases in approach behavior. An additional finding was that positive scenes elicited by evaluative words (e.g., beautiful) were more successful than serene relaxing scenes elicited by passive words (e.g., calm).

Recently Hekmat, Deal, and Lubitz (1985) further extended these procedures by adding positive self-instructions to positive imagery and image-induced relaxation. Speech-anxious clients were first asked to visualize an anxiety-provoking scene (someone

leaving the room during his or her presentation) with concomitant negative self-statements (“I am driving people away with my lousy presentation”). After 20 seconds clients were instructed to switch off the scene and relax by imagining “having fun on a date.” They were then asked to hold onto the pleasant feelings generated by relaxation and imagine the same anxiety-provoking scene again. This time the therapist instructed clients to associate positive coping statements with the scene: “It does not matter what people do or think about what I say. I will continue with my message. Besides, the person who left probably had a doctor’s appointment.” Clients were then instructed to reinforce their covert resourceful behavior by visualizing their most preferred imagery (e.g., receiving a gift) and focusing once again on the pleasant sensation of relaxation. The results of this study are indeed very intriguing: a highly credible attention placebo treatment—introduced to subjects as a new technique called systematic ventilization—also produced significant improvements on three out of four outcome measures! Subjects in the systematic-ventilization group scored consistently better than waiting-list control subjects. Most remarkably, systematic ventilization was as effective as semantic desensitization in changing affective evaluations, self-reported anxiety, and self-confidence. Only on the behavioral observation measure was the real treatment group superior to the placebo and waiting-list control groups. These findings lend some support to a criticism by Sappington (1975), whose data suggest that positive results in semantic desensitization studies might, to some extent, be due to the demand characteristics of the situation. Nonetheless, Hekmat’s treatment is quite impressive and reflects the complexities of human behavior: (a) The scenes and words were individually tailored to each client and highly credible; and (b) the treatment was far from being a “mechanistic conditioning” procedure and involved a large number of verbal, imagery, and emotional repertoires. Some theoretical concepts and labels, however, are used in a rather loose fashion and seem to be more metaphors and descriptions rather than explanations of the principles underlying the various treatment techniques. Although this may still be heuristically useful, the exact conceptual status of the multitude of treatment steps and their interrelatedness needs to be more clearly specified. In any case, Hekmat’s approach of deriving complex clinical procedures from a comprehensive theoretical framework is commendable and a step in the right direction toward bridging the gap between basic psychological theory and clinical practice.

I have recently been involved in two studies investigating the effects of language conditioning on conditioned anxiety responses. The first experiment (Eifert, 1984a) compared the effects of positive and negative self-verbalizations on the physiological, subjective-evaluative, and behavioral aspects of classically conditioned fear responses to slides of snakes. During language conditioning, these slides were paired with statements of either positive or negative affective meaning (higher-order UCS). Participants repeated these statements subvocally while the slides were on. The results showed a complete and rapid extinction of the physiological response (GSR) in groups with positive verbalizations, whereas negative statements impeded extinction. However, when subjects could observe a living snake prior to language conditioning, negative verbalizations were unable to hamper GSR extinction. The affective evaluation of snakes on semantic differential rating scales improved in all groups with positive verbalizations and deteriorated in groups with negative statements. Although

subjects in all groups with positive statements exhibited more approach behavior, this trend was not statistically significant.

In a procedurally similar second experiment (Eifert & Schermelleh, 1985), we specifically compared the conditioning effects of positive statements referring either to positive features of the phobic stimuli (snakes and rabbits) such as their beautiful skin color, or describing approach responses (e.g., "I may soon touch the animal"). As was predicted from A-R-D theory, stimulus-referent statements, aimed at the verbal-emotional repertoire, facilitated extinction of the physiological response more than response-referent verbalizations. Although snakes were consistently rated more negatively than rabbits, their affective ratings improved significantly following language conditioning with both types of verbalizations. As response-referent statements purport directly to the verbal-motor repertoire, we expected they might lead to higher approach behavior than stimulus-referent statements. However, this was not the case. It is quite conceivable that language conditioning may only affect responses to words, images, and symbols of the phobic animals—second signaling system abstractions—rather than responses to the real animals themselves; hence our failure to obtain significant improvements in approach behavior.

Both studies demonstrated that emotionally relevant language stimuli can affect the extinction of a conditioned physiological response. These results also qualify an argument frequently raised by proponents of preparedness theory (cf. Öhman, 1979) that conditioned physiological responses to fear-relevant stimuli (such as snakes) resist cognitive manipulation once they are acquired. According to social-behaviorist theory, cognitive stimuli will have an impact on existing conditioned responses if they are emotive language stimuli—and this is what we found. Furthermore, fear-relevant stimuli, such as snakes, could have more negative evaluative strength because of previous cultural learning rather than because of some "biological hard-wiring." For instance, prior to our experiment, subjects were probably exposed to many language and vicarious conditioning trials in which the word or the sight of a snake was paired with negative affective verbal stimuli, such as *ugly*, *slimy*, and dangerous situations or scenes in movies and books. On the other hand, rabbits would have been more frequently associated with positive labels like *cute* and *cuddly*. In other words, fear-relevant stimuli become more salient than other stimuli through their associations with aversive physical, verbal, and symbolic stimuli in the environment (cf. Burgess, Jones, Robertson, Radcliffe, & Emerson, 1981). This is reflected in faster response acquisition and higher resistance to extinction. Finally, it should be noted that such cultural preconditioning does not render language conditioning interventions useless or ineffective. However, it makes counterconditioning more difficult (cf. Tryon & Briones, 1985), and it means that a large number of conditioning trials over long periods of time may be necessary to overcome strong preexisting affective responses.

LANGUAGE CONDITIONING IN COGNITIVE-BEHAVIORAL INTERVENTIONS

The emergence of cognitive-behavioral interventions has divided the field of behavior therapy like no other development in its relatively short history. There is still considerable disagreement as to whether these interventions are really new and/

or effective (Latimer & Sweet, 1984). However, some of the techniques employed by cognitive therapists (e.g., cognitive restructuring) have not been traditionally used by behavior therapists, or if they have, behavior therapists did not write about them. Furthermore, recent meta-analyses (e.g., Dush, Hirt, & Schroeder, 1983) and an examination of the extensive research published in journals such as *Cognitive Therapy and Research* suggest that the effectiveness of these procedures, for at least some problems (e.g., anxiety, depression) some of the time, cannot reasonably be questioned anymore. On the other hand, there are still many unresolved conceptual problems: "What appears to be needed, in addition to carefully conceived and carried out treatment outcome research, is the development of conceptual models . . . and theoretical systems . . . to guide cognitive-behavioral procedures" (Kendall, 1984, p. 121). It is my contention that the vast amount of research on language conditioning and its integration into a social-behaviorist framework could provide a theoretical foundation and framework for cognitive-behavior therapy. Social behaviorism provides the best account available of the basic mechanisms—based in extensive research—that underlie cognitive-behavioral and verbal psychotherapy methods. Moreover, its level-by-level analysis closely links interactions between affect, behavior, and cognition to the basic learning principles. That linkage could make it possible to dispense with the cognitive versus behavioral schism that has arisen in behavior therapy. Whereas Ellis (1983) acknowledged the usefulness of verbal conditioning and social reinforcement, other proponents of cognitive-behavior therapy have increasingly distanced themselves from conditioning theories, declaring them "moribund" (Meichenbaum & Cameron, 1982). They would also probably object to cognitive interventions being interpreted within a conditioning framework. I will attempt to show, however, that this is not only possible, but it could also help to clarify the conceptual status of these interventions and provide a rapprochement between proponents of cognitive and more traditional behavior therapy (cf. Eifert, 1984b).

Meichenbaum (1977) pointed out that the various cognitive-behavioral treatments differ in terms of the relative emphasis placed on a formal logical analysis, the directiveness and forcefulness with which the therapeutic rationale and procedures are presented, and the relative reliance on adjunctive behavioral procedures. Social-behaviorist theory could help to clarify the conceptual nature of such differences. Staats (1968) suggested that a person's language repertoires (particularly labeling and reasoning) are learned and often constitute what is considered cognition: his or her reasoning, problem-solving, planning, hypothesizing, and so on. These activities involve interactions between several language repertoires and other subrepertoires. Frequently these subrepertoires are interactions between the basic behavioral repertoires as shown in Figure 1.

Cognitive-behavioral interventions differ in the repertoires that they address and utilize in order to produce change. Purely language-based interventions, such as semantic conditioning, work largely through the verbal-emotional repertoire. Words or sentences elicit an emotional response in the client, thereby changing the affective response to the stimulus or event that the verbal stimuli refer to. At other times, words themselves do not elicit the emotional response directly, but are merely used to produce a pleasant image that then elicits an emotional response (cf. Hekmat, 1977). Any behavioral changes that follow from such interventions are believed to

TABLE 1. Some Interactions between Personality Repertoires in Subrepertoires Frequently Used in Language Conditioning Interventions

Basic personality systems			
Emotional-motivational	Language-cognitive	Instrumental-motor	
Emotional	Verbal		
Emotional	Verbal	Image	
Emotional	Verbal		
	Verbal		Motor
		Image	Motor

be mediated by a change in the affective response to the signified stimuli. Interestingly, semantic conditioning has never been regarded as a cognitive intervention even though it relies entirely on language and other symbolic activities. One can only speculate that this is because semantic conditioning is based on a conditioning framework, which for those who accept the cognition versus conditioning schism, automatically means that it cannot be cognitive.

Self-Instructional Training. The work of Luria (1961) and Staats (1963, 1968) has been further developed and successfully translated into specific self-instructional training programs to treat a large variety of clinical problems, such as anxiety and anger (Meichenbaum, 1977) as well as to change the behavior of impulsive children (Kendall & Braswell, 1985). These programs and other cognitive-behavioral techniques usually combine a number of performance-based and verbal-symbolic techniques that can be analyzed within social-behaviorist theory (cf. Eifert, in press; Martin & Levey, 1985; Staats & Heiby, 1985). Frequently a reasoning or problem solving act will include (a) labeling of the event or situation, (b) sequences of verbal responses made to the labels, and (c) some final act elicited by the person's verbal processes. A close analysis of self-instructional training reveals that these three stages are integral parts of this technique. Clients are taught to emit self-statements that are incompatible with, and opposite in emotional content to, the negative self-statements they have employed previously. As the emotional value of labels affects the client's reasoning, problem-solving, and overt behavior, the therapist assists the client in changing these labels. The principles of semantic counterconditioning and verbal reconditioning are employed in this stage. For instance, clients learn to label physiological arousal as a cue to employ coping skills rather than a sign of an imminent panic attack. In addition, clients are instructed to use the verbal-motor repertoire to direct their overt behavior and thoughts in such a way that they can better cope with difficult situations (e.g., "one step at a time"; "just think about what you have to do"). Finally, reinforcing self-statements ("you did it, it worked") are used to maintain the newly acquired coping skills. The results of clinical outcome studies as well as the standard procedure and procedural variations of self-instructional training have been described in great detail by Meichenbaum (1977); the interested reader is referred to this inspiring book and a meta-analysis by Dush, Hirt, and Schroeder (1983) on the effectiveness of self-instructional training.

The treatment of delusional and other "crazy" talk sometimes engaged in by schizophrenics was one of the first applications of verbal conditioning and self-instructional training programs. Several early studies (e.g., Krasner, 1958; Meichenbaum, 1969; Nydegger, 1972) not only demonstrated that "sick talk" and bizarre verbalizations could be eliminated or at least greatly reduced by verbal conditioning methods, but that social reinforcement (e.g., attention) also serves to maintain such inappropriate behaviors. Therefore I will briefly review more recent findings in that area. Interestingly, it was the serendipitous finding of the spontaneous use of self-instructions by schizophrenics during a verbal conditioning treatment (Meichenbaum, 1969) that initiated a research program on "training schizophrenics to talk to themselves" (Meichenbaum & Cameron, 1973). Self-instructional training was used to teach schizophrenic patients to use their overt speech to monitor their own behavior and thinking. They were also trained to become sensitive to interpersonal signals from others that indicated that they were emitting bizarre, incoherent, or irrelevant behaviors and statements. Both the interpersonal observations and the self-monitoring provided cues to emit a set of self-controlling instructions (e.g., "I must be relevant and coherent to make myself understood"). Patients in this self-instructional training improved, and continued to improve, significantly more than patients treated with straightforward verbal operant conditioning. Although in this study reactive paranoid and chronic process schizophrenics were treated successfully, there is empirical evidence that whereas verbal conditioning is effective for reactive schizophrenic clients, it is less effective or even ineffective for chronic process schizophrenic clients. A study by Pansa (1979) suggests that the degree of premorbid social responsiveness may be responsible for this finding. Two other studies (Caulfield & Martin, 1976; Miller & Drennen, 1970) demonstrated, however, that even chronic schizophrenic patients can benefit from verbal conditioning if verbal and social reinforcements are at first paired with primary reinforcers.

The clinical evidence on the effectiveness of verbal conditioning and self-instructional training with severely disturbed psychiatric patients is very impressive. Even though psychoactive drugs have helped to control some of the bizarre verbal and other behaviors, no drug can (yet) teach new skills. If we are serious about preventing relapses, these patients should be taught cognitive and social skills in self-instructional programs to maintain any drug or otherwise induced changes.

Cognitive Restructuring. In cognitive restructuring client labeling and reasoning repertoires are changed through direct communication interactions with the therapist. The therapist challenges the appropriateness and empirical validity of client beliefs, appraisals, interpretations of situations and events, as well as the way clients categorize and process information ("cognitive styles").

One of the most frequent applications of cognitive therapy has been in the area of depression (e.g., Beck, 1976). Verbal and semantic conditioning studies have attempted and succeeded in achieving similar effects. Coons, McEachern, and Annis (1973) were able to increase self-acceptance responses in a group of mental hospital patients by means of verbal conditioning. Treatment effects also generalized and improved clients' general attitudes toward themselves and other persons. In a very interesting study (Sappington *et al.*, 1982), the concept *myself* was the target of semantic

imagery conditioning even though the subjects were not depressive but phobic. During five sessions of semantic conditioning the stimulus *self* was paired with “powerful images” elicited by words scoring high on the potency dimension of the semantic differential. The results show that subjects’ behavior test performance significantly improved compared to a pseudoconditioning control group. This study is particularly interesting because it suggests that a person’s level of perceived self-efficacy—which Bandura (1984) regards as the most important mediator of therapeutic change—can actually be modified by semantic conditioning procedures. Sappington and his colleagues raise a number of important issues in discussing these and their earlier findings, concluding that

the intriguing possibility exists that procedures such as semantic conditioning, systematic desensitization, implosion, etc., work because they change beliefs (specifically emotionally based assessments) and yet they might change beliefs through an associationistic process. (p. 322)

In other words, therapists could explicitly use language conditioning techniques and extinguish their clients’ negative statements about themselves, while reinforcing more positive statements. As clients learn to associate themselves more frequently with positive statements and labels, their self-perception and self-concept should gradually become more positive, too, increasing their level of perceived resourcefulness (cf. Rosenbaum, 1983).

EFFECTS OF LANGUAGE CONDITIONING ON VERBAL, AFFECTIVE, AND BEHAVIORAL CHANGES

It is of great interest to behavior therapists to understand the interplay between overt language, inner speech and thought, emotional experience, and overt behavior. From a clinical perspective one of the most pertinent questions is whether changes brought about by language conditioning interventions also generalize to the client’s behavior and feelings “in the real world” outside the therapist’s office. In other words, how broad and generalized are the effects of language conditioning? These questions will be examined in the following sections.

MODIFYING CLIENTS’ VERBAL BEHAVIOR IN THERAPY: AN ANALOGUE TO PSYCHOTHERAPY?

As early as 1958 Krasner pointed out that one of the most fruitful and immediate applications of verbal conditioning lies in the modification of clients’ verbal behavior during therapy sessions. Krasner argued that essentially all psychotherapy is to some extent directive in nature and could be viewed in the context of verbal conditioning. This notion challenged earlier views that some forms of therapy can be nondirective. Extensive subsequent research by Truax (1966) and Martin (1975) confirmed that even client-centered therapists respond selectively to different verbalizations of their clients. They tend to reinforce affective self-references by reflecting and rephrasing their emotional content, whereas they largely ignore (extinguish) reports of more

factual-type "objective" material. There is indeed ample empirical evidence showing that reflection of feeling content as well as the straightforward reinforcement of affective self-references increases the frequency of such statements. This has been shown for quite diverse populations ranging from college students (Highlen & Nicholas, 1978) and patients in mental hospitals (Coons, 1972) to imprisoned drug addicts (Hafner & Linkenhoker, 1974).

On the other hand, Krasner (1982) cautioned that verbal conditioning and psychotherapy are not the same identical process, nor is one an analogue of the other. Some verbal conditioning, however, takes place in any therapy, and the relationship variables involved in therapy cannot and should not be eliminated from the study of verbal conditioning. Similarly, Lilliston (1972) argued that any therapy involves an interaction of therapist, client, problem, and treatment strategy. Verbal conditioning relates to only one of two aspects of selecting a treatment strategy: it is one possible option for the therapist in treating a particular problem, but it does not refer to the decision as to what content should be dealt with. For instance, studies testifying that the frequency of affective self-references can be increased by verbal conditioning say nothing about the desirability of making affective statements a treatment target. The choice of an appropriate target behavior should be based on a thorough psychological assessment and functional analysis of the problem.

LANGUAGE CONDITIONING AND EMOTIONAL AROUSAL

Cognitive-behavior therapists in particular have repeatedly emphasized that emotional arousal is mediated by negative idiosyncratic self-statements. Studies on semantic conditioning and generalization indeed support the notion that through conditioning certain word stimuli come to elicit physiological arousal. Similarly, Bandura (1984) pointed out that emotional disorders are, in large part, self-generated and the product of self-referent thoughts and verbalizations rather than automatically evoked by conditioned stimuli. As will be discussed in the following, there is evidence that at least some of the dysphoria, physiological arousal, and anxiety experienced by individuals is self-generated and affected by the individuals' appraisal and labeling of themselves and/or the situation. The stimulus properties of self-statements and images (conditioned sensory responses) can actually elicit affective responses in the individual and direct approach or avoidance behavior (Staats, 1975). This qualifies contentions that conditioning theorists treat thoughts merely as a by-product of conditioned responses (e.g., Bandura, 1984). Martin and Levey (1985), pointing out that such statements reflect an anachronistic view of contemporary conditioning theories, discuss the important role of evaluative cognitions and how these can be modified through conditioning interventions. Very important functions of language reside in the affective qualities of words. As noted earlier, thoughts and self-statements can come to function as (higher-order) UCS capable of generating dysphoria and depression (Staats & Heiby, 1985) as well as anxiety (Eifert, in press), but they can also be used to countercondition such negative responses (Eifert, 1984a). In other words, although in some cases words and thoughts are merely descriptions and concomitants of conditioned responses, in many other cases they *are* conditioned stimuli and responses. Considering the potential conditioning trials that are available

in a person's lifetime, intense emotional reactions can easily be accounted for on the basis of language or other forms of symbolic conditioning. Acknowledging and explaining this important role of language and symbolic stimuli is one of the most valuable contributions of social behaviorism to the field of behavior therapy.

However, the relationship between self-statements and physiological arousal is quite complex and neither as straightforward as cognitive-behavior therapists assume nor as automatic as one might be tempted to conclude from some language conditioning studies. For example, the conditioning of highly personally relevant negative self-verbalizations may produce considerable physiological arousal: Master and Gershman (1983) found that in high relevance situations, "irrational" verbalizations led to greater physiological arousal than rational statements. On the other hand, Rogers and Craighead (1977) showed that neither minimally nor highly discrepant statements, but those that were moderately discrepant from the subjects' current beliefs, produced the highest physiological arousal. They also found that positive and negative statements produced similar levels of physiological arousal.

LANGUAGE CONDITIONING AND BEHAVIORAL CHANGES

Before discussing the effects of language conditioning on actual overt behavior, it is important to review some of the basic principles regarding the interplay of verbal and motor behavior. In his work with children and brain-damaged persons, Luria (1961) focused on the important behavior-directive function of language. Children acquire this ability of language to control overt behavior in three stages: initially, verbal cues from adults are most pertinent; from that children learn to direct their behavior by talking to themselves aloud; and finally, this overt speech turns into subvocal speech and thoughts. Staats (1968) also elaborated on the behavior-directive function of a large number of verbal stimuli organized in the verbal-motor repertoire.

In a similar vein, cognitive-behavior therapists have consistently argued that changing labeling and reasoning will also change behavior. However, the exact relationships between different types of self-statements and overt behavior are not clearcut as one might think. For instance, positive statements do not always lead to approach behavior and negative statements do not always lead to avoidance behavior. We have recently examined the relationships between different types of spontaneous self-verbalizations and overt behavior to a fear-provoking animal (Eifert & Lauterbach, 1987). As in a previous study (Eifert & Schermelleh, 1985) we found that approach behavior was consistently associated with a greater number of more of positive compared to negative statements. Subjects approached fear stimuli as long as they had enough positive statements to compete with the negative ones. These results suggest that, once negative statements have been detected, the therapist should concentrate on teaching clients the use of positive statements to dispute existing negative thoughts. Thereby positive self-verbalizations will be more frequently associated with stressful situations, and as Rosenbaum (1983) pointed out, gradually become part of a larger repertoire of "learned resourcefulness." Positive coping self-statements are considered an integral part of this learned cognitive-behavioral repertoire. They enable individuals to control and minimize the undesirable and interfering effects of negative thoughts and statements on the smooth execution of a target behavior.

Basic experiments with nonclinical populations and target behaviors (summarized in Staats, 1975) have repeatedly been able to demonstrate the effects of language conditioning on the control of overt behavior. Clinical studies investigating the effects of language conditioning on overt behavior have produced somewhat more mixed results—particularly those that actually tried to countercondition already existing strong affective responses. Some studies (e.g., Hekmat, 1972, 1977) were able to demonstrate that overt behavior changed as a consequence of language conditioning. However, other studies (e.g., Weiss & Evans, 1978) found only marginally significant ($p < 0.10$) within-group changes in the semantic conditioning groups but no significant differences between the semantic conditioning groups and the control groups! Similarly, our own studies failed to show significant changes, although the trend was in the predicted direction. Unfortunately, other interesting studies did not employ behavioral measures at all. For instance, following quite an elaborate semantic conditioning procedure, Tryon and Briones (1985) obtained significant counterconditioning effects on semantic differential ratings of various sexual activities. Even though a 6.3 rating on a 7-point scale is a statistically significant decrease from a preexperimental level of 6.8, it is still a very negative rating of sexual activities. Such relatively small affective-attitudinal changes make it seem unlikely that any substantial change in overt sexual behavior would have occurred even if it had been measured.

Meichenbaum (1977) reviewed a number of studies where verbal conditioning changed clients' verbal but not their motor behavior. As human learning is cumulative-hierarchical, it is important to ensure that the prerequisite motor behavior that language is supposed to control is already in the clients' repertoire, particularly when treating children and severely disturbed adults. Furthermore, Meichenbaum pointed out that clients should be encouraged to verbalize prior to acting in order to facilitate the regulatory function of language. Such "say-do" training capitalizes on the fact that verbal cues are more readily available and versatile discriminative stimuli than nonverbal cues and thus more likely to prompt rehearsal. Reinforcement of verbal behavior alone may result in only slight increases in corresponding nonverbal behavior, whereas reinforcement of a correspondence between verbal and nonverbal behavior will lead to an increased correspondence.

CONCLUSIONS

It must be clearly recognized that verbal and symbolic stimuli are only one source of behavioral regulation. Rachman (1981) referred to the somewhat limited potential of language in changing existing strong affective responses, even though language can be quite powerful in inducing or establishing affective responses. This is an interesting and very important distinction with implications that have not yet been fully investigated. It may explain why in some cases the effects of language conditioning generalize from one response system to another and why this does not happen in other cases. Treatment programs that were aimed at either establishing verbal response classes or strengthening existing ones have in fact often obtained a generalization from verbal changes to behavioral and affective changes (e.g., Coons, 1972). On the other hand, whilst studies attempting to change or countercondition existing affective responses or overlearned habits (e.g., addictions) may succeed in

obtaining some degree of affective and cognitive changes, particularly on self-report measures, these changes may not be strong enough to produce lasting behavioral changes as well. Such problems may have to be treated by verbal conditioning in conjunction with performance-based interventions.

A similar position, described as a "therapeutic paradox," has been adopted by most proponents of cognitive-behavior therapies (e.g., Ellis, 1979; Meichenbaum & Cameron, 1982). These authors have repeatedly emphasized that performance-based treatment methods are significantly more effective in producing behavior change than methods that rely solely on verbal or imaginal procedures. Ellis (1979) admitted that pure cognitive restructuring works relative poorly when trying to change strong emotional reactions such as those involved in complex disorders like agoraphobia. Individuals must act against their irrational beliefs if they really want to overcome them. Nevertheless, studies on a variety of clinical problems (e.g., Eifert, 1984b; Marshall, 1985) have shown that the maintenance of behavior change is enhanced by combining performance-based procedures with verbal techniques, such as self-instructional training. It should be noted that language is probably also part of behavior-based treatments: clients are presumably thinking in words, even if they are not consciously using verbal regulation techniques.

We must also be careful not to generalize from the successful use of a particular treatment technique to statements about the origin of a clinical problem. For instance, even if language conditioning works to alleviate an emotional problem, we cannot conclude that the problem was acquired in the same way or, alternatively, that a successful behavior-based treatment proves that language conditioning had no part in the origin of the problem.

Finally, research on the interface between affect, behavior, and cognition (Izard, Kagan, & Zajonc, 1984) indicates that the functional relationships between central affective responses, peripheral physiological arousal, overt behavior, and the A-R-D functions of language are very complex and, indeed, far from fully understood. For instance, whether or not a negative evaluation of an object leads to avoidance behavior depends on a multitude of factors. The individual's learned labeling repertoire is only one, albeit an important one of these factors. As noted above, we found in our studies that at times fear stimuli were approached in the presence of very negative self-statements. This does not imply that there may not be a general tendency or predisposition to act in a way that corresponds with one's verbal and affective responses.

There are also a number of measurement problems that make it difficult to assess whether self-verbalizations are actually controlling overt behavior or whether they are merely a verbal commentary to individual emotional experience and behavior. In fact, we have to consider the possibility that they can be both. Beck and Emery (1985) have recently proposed an interesting modification to the notion that language or symbolic activities actually cause depression or anxiety disorders. They argue that panic attacks are precipitated by inappropriate statements, thoughts, and images that signal danger:

The crucial element in anxiety states, thus, is a cognitive process that may take the form of an automatic thought or image that appears rapidly, *as if by reflex*, after the initial stimulus (e.g., shortness of breath), that seems plausible, and that is followed by a wave of anxiety. (p. 5, italics added)

The reference to the reflex-type character of these thoughts and images is very interesting because it indicates that classical conditioning may be involved in the formation of such cognitive reflexes. These views are quite congruent with the social-behaviorist model: verbal or symbolic stimuli, which may be elicited by some physiological change or environmental event, can actually precipitate and intensify an emotional response. The structural organization of language, affective, and motor responses in overlapping basic behavioral repertoires or personality systems (see Figure 1) would, in any case, caution against one-sided cause-effect conclusions.

Taken together, the previously cited findings and arguments strongly suggest that in some cases it is indeed possible for psychotherapy to take place on a solely verbal level: "Deficit behavior, inappropriate behaviors, stimulus control, the reinforcer system should all be accessible to change through verbal means" (Staats, 1972, p. 167). On the other hand, language-conditioning procedures in clinical practice have their limitations in that verbal-cognitive changes do not always generalize to changes in behavior. Yet recognizing and specifying those limits is far from saying that language conditioning is useless and we should abandon our research and clinical efforts in that area. We should merely be alerted to the fact that in many cases it will be necessary to combine language conditioning procedures with performance-based techniques.

VARIABLES INFLUENCING THE EFFECTIVENESS OF LANGUAGE CONDITIONING

Extensive experimental and clinical research has identified a number of variables that influence the effectiveness of language conditioning. Knowing and incorporating these findings into clinical work should be useful for behavior therapists and enhance the efficacy of their verbal interactions with clients. These variables include: (a) personality characteristics and interpersonal skills of the therapist; (b) personality differences between clients as well as the evaluative strength and credibility of the verbal material presented by the therapist; (c) client awareness of the conditioning contingencies and the problem of client countercontrol and resistance.

PERSONALITY CHARACTERISTICS AND INTERPERSONAL SKILLS OF THE THERAPIST

It has almost become a truism that reinforcement does not work in an automatic "unconscious" fashion and its effectiveness is to no small extent determined by whom and how it is delivered. Wilson and Evans (1977) noted that the importance of a good client-therapist relationship and the impact of the therapist's personality on client feelings and behavior has only gradually been realized in behavior therapy, although it has always been a central tenet of other therapies such as client-centered therapy. Namenek and Schuldt (1971) found that therapists scoring high on empathy, warmth, and genuineness yielded better verbal conditioning effects than therapists scoring low on these characteristics. Clients of such therapists learn to communicate and understand their emotional states more effectively and accurately, pay more

attention to their therapist, and are more likely to accept their therapist as a model (cf. Wilson & Evans, 1977). These changes would certainly be considered positive regardless of whether therapists label themselves as client-centered, rational-emotive, or behavioral. Therapists who score high on the "client-centered triad" are also liked more by their clients and are more effective in verbally conditioning their clients (cf. Brown, Helm, & Tedeschi, 1973). Wilson and Evans (1977) pointed out that social-behaviorist theory would indeed predict these findings: as the emotional and reinforcing properties of a stimulus are interrelated, it follows that the more positive an emotional response the therapist is able to evoke in the client, the greater will be the therapist's reinforcing value and his or her ability to direct the client's overt behavior.

PERSONALITY CHARACTERISTICS OF CLIENTS

In his presidential address to the Association for the Advancement of Behavior Therapy, Ross (1985) criticized behavior therapists for having neglected research on individual differences and personality characteristics (see also Eysenck, 1982). Differences in personality repertoires and other interindividual differences have indeed a profound impact on the conditionability of clients and hence mediate the effectiveness of language conditioning.

Personality Differences between Clients. What a person learns in a language conditioning experience, or any other given situation, not only depends on the stimulus characteristics of the situation, but also on the unique make-up of that person's learned personality repertoires. In a series of experiments, Staats (1980) demonstrated that individual differences in the emotional-motivational personality system determined how emotional stimuli were perceived, experienced, evaluated, and which new responses were learned. For instance, depending on subjects' preexisting particular value systems, they learned quite opposite emotional and affective responses although they were subjected to the same language conditioning experience.

There are also a number of more specific personality factors that mediate the effects of language conditioning (Eysenck, 1960, 1982). It has been repeatedly found (e.g., Gupta, 1973) that verbal conditioning effects are stronger and more resistant to extinction in introverted than in extraverted persons, and that stimulants (e.g., dextrine) facilitate the conditioning process whereas depressants (e.g., phenobarbitone) retard conditioning. Eysenck (1964) maintained that delinquents are more extraverted than nondelinquents and should therefore be less amenable to social reinforcement. Johnston (1976) conducted six experiments and found that whereas the delinquents in her sample were not more extraverted than nondelinquents, they were generally less responsive to social reinforcement and punishment. Jolley and Spielberger (1973) found that anxiety and locus of control interact in their effects on verbal conditionability: high-anxiety externals were more responsive to social reinforcement than high-anxiety internals, and low-anxiety internals were more responsive than low-anxiety externals. Highlen and Nicholas (1978) showed that self-referenced affective statements could be increased for both internal and external locus of control subjects, provided they were given specific instructions prior to conditioning. It has been found that conditioning effects are stronger for subjects who have a

high need for social approval and who are highly susceptible to hypnosis (King & McDonald, 1976). Furthermore, persons of low intelligence conditioned better than average and bright persons (Mohan & Dharmani, 1976). However, intelligence may interact with the ability to detect conditioning contingencies. Javierto (1971) found that subjects who were aware of the experimental contingencies conditioned sooner, but they were also the more intelligent subjects.

Evaluative Strength of Verbal Stimuli. The evaluative strength of verbal stimuli refers to the intensity of the emotional response elicited by different verbal stimuli. Several studies (see Staats, 1968) suggest that the classical conditioning of meaning establishes the reward value of verbal reinforcers. Words that elicit affective responses will also function as reinforcers. Hekmat and Lee (1970) demonstrated that there is a relationship between the strength of the evaluative meaning of words and their reinforcement value: analysis of semantic differential ratings of the verbal reinforcers indicated, not surprisingly, that "wonderful" was rated more positively than "good," which in turn was rated more positively than "mmm-hmm." More importantly, the efficacy of these words in increasing affective self-references was in the same order. The results also showed that how positively a person rated these words clearly determined the strength of conditioning in the individual case. One would assume that the language habits of the reinforcing person are also influential in determining evaluative strength: "Wonderful" from someone who habitually speaks extravagantly may not have the same reinforcement value as "good" from someone who is sparing of praise. Similarly, it should be noted though that satiation and habituation effects may occur. Hekmat (1974) found that if the same reinforcers were used too often, their subsequent effectiveness attenuated. Therapists are therefore advised to employ a variety of verbal reinforcers with high evaluative strength. The evaluative strength can be determined for each client individually by asking him or her to rate a list of such words on semantic differential scales at the beginning of therapy. Furthermore, in regard to reinforcement schedules the same principles established in other conditioning situations seem to apply in verbal conditioning situations as well: continuous reinforcement results in high response frequencies, but intermittently reinforced verbal responses are more difficult to extinguish than those acquired under continuous reinforcement.

Credibility of Verbal Material and Individualization of Treatment. This issue refers to the potential problems arising from an incongruence between words or statements used in language conditioning therapies and the client's existing evaluations and beliefs. In other words, how much do clients have to agree with the verbal stimuli that are used to condition them?

Evidence from a study by Lilliston (1972) suggests that subjects' preexisting values mediate the effectiveness of verbal conditioning procedures: subjects conditioned to words consistent with their value orientations produced a steeper acquisition curve than subjects conditioned to words that were inconsistent with their preexperimental values. Relating his results to verbal conditioning in therapy, Lilliston concluded that therapists should consider personality and value differences between clients as important determinants of conditionability. Similarly, Laungani (1970) demonstrated in a series of experiments with school children that previously learned

verbal habits exercised greater influence on the children's choice of verbal responses than the reinforcement presented by the experimenter.

Discrepancies between clients' existing verbal-emotional repertoires and labels employed by the therapist is a problem that may occur in any language conditioning procedure. If these discrepancies are large enough, the client's past learning will result in verbal behavior that may actually counter the conditioning provided by the therapist. For instance, a depressed person is likely to label him or herself negatively, just as a phobic person is likely to label a potentially fear-provoking stimulus in a negative way. Such negative labeling may counter the therapist's positive conditioning wherein he or she tries to associate more positive labels with the person or the phobic object. Staats (1975) suggested that the same conditioning results whether the verbal stimuli are produced by oneself (the client) or by another person (the therapist). Yet it is not quite clear what will happen if these two sources are too discrepant or even opposite and mutually exclusive. Is the therapist's conditioning only effective if he or she can somehow prevent the client from emitting "counter-therapeutic" language/thought stimuli? Tentative answers to these crucial questions come from social psychological and clinical research. Craighead and Craighead (1980) found that maximal attitude change will occur at moderate levels of discrepancy from existing beliefs. Such information is likely to be assimilated rather than rejected by the client. It follows that positive self-statements employed in therapy should not be too positive—particularly not at the beginning of therapy. Such statements are likely to be contrasted with existing attitudes and may be perceived as too discrepant from existing beliefs, resulting in no change at all. Craighead and Craighead suggested that a gradual shaping program, similar to a desensitization approach, might be most effective, starting with self-verbalizations that are only moderately discrepant. Marshall (1985) examined the timing of self-instructional training during an exposure program for acrophobic clients. He pointed out that relabeling interventions should only be employed after some initial anxiety reduction has occurred. In this case, positive labels and appeasing statements are less likely to be rejected by a client because at that stage such statements will be more in line with what he or she actually experiences.

Words to be used in therapy situations should therefore have a clear and obvious thematic-semantic relationship to the phobic object. Evans and Weiss (1978) demonstrated that such words are, indeed, more effective in language conditioning procedures. This is supported by social psychological research (e.g., Craighead & Craighead, 1980) showing that attitudes and behavior can be more effectively changed if the person takes an active role in generating the content of the counterattitudinal message. Therefore clients should actively participate in the generation and selection of UCS words and in the formulation of sentences to be used in conditioning treatment programs. It is very encouraging to note that more recent studies (e.g., Hekmat *et al.*, 1985) have adopted this approach. Such active participation of clients not only increases the credibility, individuality, and effectiveness of treatment, but is also in tune with the prevailing zeitgeist in behavior therapy. Clients are no longer seen as subjects to be conditioned but as persons who are actively involved in the resolution of their problems. This should also refute the common stereotype and criticism that treatment based on conditioning principles is by its very nature necessarily mechanistic and impersonal.

CLIENT KNOWLEDGE OF CONDITIONING CONTINGENCIES AND COUNTERCONTROL

Awareness and Conditionability. Almost no other topic in the area of verbal conditioning has generated so many studies and controversial findings as has research on the role of client awareness or knowledge of the experimental contingencies. Some researchers (cf. Ladouceur & Mercier, 1984) have consistently found that unless clients are aware and able to verbalize the S-R contingencies, conditioning does not occur. Such findings have been used to argue for a cognitive interpretation of the effects typically found in conditioning studies (cf. Bandura, 1969, 1984). Conversely, other researchers have shown that contingency awareness is not necessary for successful conditioning (e.g., Burgess & Linder, 1970). However, research on the role of awareness in language conditioning is surrounded by a plethora of thorny methodological and conceptual problems. Ladouceur and Mercier (1984) rightfully pointed out that considering awareness-unawareness as a dichotomous rather than a continuous variable is a major source of confusion. For instance, there is the possibility that subjects may be aware of the fact that they are being conditioned, but they may be unable to verbalize the contingencies or remember them when they are interviewed about them following the experiment. Furthermore, people can respond to a stimulus in the absence of the ability to report verbally its existence.

The methods of assessing awareness have also been criticized (Martin & Levey, 1978; Staats, 1975). For instance, some of the questionnaires or standardized interviews may not just measure awareness but actually produce awareness and alert people to the contingencies—which would not have happened if they had not been asked in the first place. I do not intend to reopen this controversy and/or take a side for either position. Bandura's (1969) conclusion is still a fair summary of the majority of findings:

The overall evidence would seem to indicate that learning can take place without awareness, albeit at a slow rate, but that symbolic representation of response-reinforcement contingencies can markedly accelerate appropriate responsiveness. (p. 577)

In a recent review on the role of awareness in behavior therapy, Ladouceur and Mercier (1984) came to a very similar conclusion. There is little doubt that there are definitely cases of learning without awareness but in some—albeit not all—instances such learning may be slower, weaker, and easier to extinguish.

Awareness and Countercontrol. Ladouceur and Mercier (1984) also postulated that awareness of conditioning contingencies parallels and facilitates behavioral change because it accelerates performance change. This conclusion is of particular relevance for behavior therapy. It must be qualified, however, because the facilitative influence of contingency awareness will partly depend on client willingness or preparedness to be conditioned. For instance, Rhodes, Shames, and Egolf (1971) conducted a study in which two groups of stutterers were verbally rewarded for talking about “desirable” and punished for talking about “undesirable” language themes. One group was informed of the reinforcement contingencies and the other group was not informed. Both aware and unaware subjects exhibited increments in their language responses following positive reinforcement, but only unaware subjects showed consistent decrements in stuttering frequency when they were punished for such responses by the

therapist's expressing disapproval. These interesting results suggest that contingency awareness may not be therapeutically beneficial in all cases. Although it can only be inferred from the results, it appears that the aware subjects responded to, or accepted the positive reinforcement contingency, but they rejected and acted against the punishment contingency. One is almost tempted to conclude that they simply refused to change their behavior. On the other hand, the behavior of subjects who did not suspect that they were conditioned did change in both directions. Hekmat and Theiss (1971) also found that high self-actualizing persons, measured by the Personal Orientation Inventory, resisted social reinforcement during interview sessions and were not very responsive to withdrawal of reinforcement in extinction trials. In other words, the (verbal) behavior of individuals who are strongly autonomous and independent is less likely to be influenced by environmental contingencies. An intriguing aspect of this study is that a conditioning study provided support for a basic tenet of humanistic psychologists such as Rogers and Maslow.

This problem of countercontrol and resistance to conditioning is in fact a very important issue. One can argue that clients normally seek therapeutic help because they want to change and that in most cases they will therefore not resist interventions that are designed to achieve the changes desired. Nevertheless, the problem of countercontrol has been underestimated in verbal behavior therapy, and I concur with Wilson and Evans (1977) that the therapist-client relationship needs to be conceptualized as an interacting system of mutual social influence:

Verbal conditioning studies have assumed a one-way influence process in which the therapist has predetermined which responses will be reinforced on what schedule and in what manner. . . . Whether a client emits the targeted behavior or engages in countercontrol will depend on a number of complex interacting factors, including the situational context, the nature of the incentive [and the] personal characteristics of both therapist and client. (p. 530)

CONCLUSIONS

One of the main objectives in writing this chapter was to find a balance between two rather extreme positions. On the one hand, there is the outright rejection of conditioning models (particularly by some proponents of cognitive therapy). It has to be recognized that some researchers turned away from conditioning models as a result of negative findings of studies testing conditioning-based therapies; an example is Meichenbaum's (1977) discussion of the awkward conceptual problems that ensued from the results of some studies on anxiety-relief conditioning. At the same time, though, a more philosophically based dislike of conditioning models has pervaded the cognitive-behavioral literature, which may be a reflection of a more general change in zeitgeist favoring cognitive theories (cf. Mahoney, 1977). The other extreme position is the argument that almost all therapy can take place on a solely verbal level if it only follows the principles of classical and operant conditioning.

The findings and arguments presented in this chapter may be summarized by concluding that the principles of conditioning do apply to language and symbolic stimuli. These activities are not (at least not always) so complex that they cannot

be adequately accommodated in conditioning terms and models. At the same time, qualitative differences between language and other stimuli have to be taken into account. For instance, Levis and Malloy (1982) pointed out that language and other symbolic stimuli are weaker because they are secondary or higher-order UCS: they cause no pain or tissue damage, reduce no primary drive, and so on. Furthermore, language conditioning does not occur automatically in a passive individual in a reflex-type manner, but is mediated by several variables of the person and the therapist. An important factor influencing the effectiveness of language conditioning is how well it is tuned to a client's existing verbal-emotional repertoires. In other words "faulty" statements should not be counterconditioned or replaced like a mechanic replaces faulty spark plugs (Coyne, 1982). To be maximally effective, words or sentences used in language conditioning programs should be individualized, largely self-generated by the client, credible, and optimally discrepant to existing verbal repertoires.

Although the limitations of language conditioning should be recognized, cognitive and other behavior therapists are well advised not to disregard the experimental and clinical findings on the effects of language conditioning. In more than 30 years of research it has been demonstrated that therapists are able to change their clients' verbal and at times also their nonverbal overt behavior by means of language conditioning techniques. No therapist can afford to ignore such findings and the significance of verbal conditioning effects because, as Krasner (1958, 1982) pointed out, they are part of almost every interaction between therapist and client. Some readers may shudder at the very idea of attempting to condition clients at a time when the active participation of clients is emphasized in every textbook on clinical intervention. Yet it is quite compatible to implement language conditioning programs—or be a behavior therapist for that matter—and to be client-centered at the same time. It has been shown that such programs can be individually tailored with the active involvement of the client. The client's cooperation and full awareness of the contingencies as well as a good therapeutic relationship, enhanced by a warm, empathic, and genuine therapist, are not just socially and ethically desirable, but are pertinent factors influencing the effectiveness of language conditioning.

Since the late 1970s cognitive-behavioral interventions have almost exclusively been conceptualized in some cognitive theory framework (cf. Bandura, 1984; Beck & Emery, 1985). Some readers will therefore object to analyzing them in a conditioning framework and discussing them along with semantic and covert conditioning. I have attempted to show, however, that language conditioning principles, such as verbal reinforcement, extinction, counterconditioning, and the learning of new labeling and reasoning repertoires, are integral parts of cognitive-behavioral interventions. Cognitive-behavior therapists prefer a different jargon for their theoretical concepts and treatment techniques. Analyzing verbal-cognitive interventions within the social-behaviorist paradigm, however, enables us to incorporate the findings of verbal and semantic conditioning studies systematically into behavior therapy techniques, thereby refining them and also increasing their effectiveness. Such analyses might also help overcome the unfruitful current schism between cognitive and behavioral/conditioning models and interventions. What cognitive and behavior therapists actually do when they see a client is often much more similar than what they say or think they

are doing (Mahoney, 1979). A further ground for rapprochement is the finding that the overall treatment effectiveness can often be enhanced by combining verbal-cognitive techniques with performance-based interventions. This is particularly important for the maintenance and generalization of treatment effects. Differences between the two approaches have been exaggerated, overemphasized, and are to no small extent due to differences in preferred terminology. A social-behaviorist analysis could help overcome these largely artificial differences and provide a bridge for rapprochement by reestablishing the links between basic experimental and applied clinical research.

ACKNOWLEDGMENTS

I wish to express my gratitude to Betty Drinkwater, Virginia McKendrick, and Lyn Craill for their constructive, helpful, and encouraging comments.

REFERENCES

- Bandura, A. (1969). *Principles of behavior modification*. New York: Holt, Rinehart & Winston.
- Bandura, A. (1984). Recycling misconceptions of perceived self-efficacy. *Cognitive Therapy and Research*, 8, 231–255.
- Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International University Press.
- Beck, A. T., & Emery, G., with R. L. Greenberg (1985). *Anxiety disorders and phobias—a cognitive perspective*. New York: Basic Books.
- Brown, R. C., Helm, B., & Tedeschi, J. T. (1973). Attraction and verbal conditioning. *Journal of Social Psychology*, 91, 81–85.
- Brownell, K. D. (1984). The addictive disorders. In C. M. Franks, G. T. Wilson, P. C. Kendall, & K. D. Brownell (Eds.), *Annual review of behavior therapy*, Vol. 8. New York: Guilford.
- Burgess, I. M., & Linder, M. J. (1970). The effects of a barrier upon verbal conditioning and awareness. *Journal of Behavioral Science*, 1, 69–75.
- Burgess, I. S., Jones, L. M., Robertson, S. A., Radcliffe, W. N., & Emerson, E. (1981). The degree of control exerted by phobic and non-phobic verbal stimuli over the recognition behavior of phobic and non-phobic subjects. *Behavior Research and Therapy*, 19, 233–243.
- Caulfield, J. B., & Martin, R. B. (1976). Establishment of praise as a reinforcer in chronic schizophrenics. *Journal of Consulting and Clinical Psychology*, 44, 61–67.
- Cautela, J. (1973). Covert processes and behavior modification. *Journal of Nervous and Mental Disease*, 157, 27–36.
- Coons, W. H. (1972). Psychotherapy and verbal conditioning in behavior modification. *Canadian Psychologist*, 13, 3–29.
- Coons, W. H., McEachern, D. L., & Annis, H. (1973). Self-acceptance, acceptance of others, and verbal conditioning with mental hospital patients. *Canadian Journal of Behavioral Science*, 5, 290–296.
- Coyne, J. C. (1982). A critique of cognitions as casual entities with particular reference to depression. *Cognitive Therapy and Research*, 6, 3–13.
- Craighead, L. W., & Craighead, W. E. (1980). Implications of persuasive communication research for the modification of self-statements. *Cognitive Therapy and Research*, 4, 117–134.
- Drinkwater, B. A. (1968). Generalization of responses to verbal stimuli. *Australian Journal of Psychology*, 20, 179–183.
- Drinkwater, B. A. (1972). The significance of affect in verbal learning by subnormal children. *Australian Journal of Psychology*, 24, 327–329.
- Dush, D. M., Hirt, M. L., & Schroeder, H. (1983). Self-statement modification with adults: A meta-analysis. *Psychological Bulletin*, 94, 408–422.
- Eifert, G. H. (1984a). The effects of language conditioning on various aspects of anxiety. *Behavior Research and Therapy*, 22, 13–22.

- Eifert, G. H. (1984b). Cognitive behaviour therapy: A critical evaluation of its theoretical-empirical bases and therapeutic efficacy. *Australian Psychologist*, *19*, 179–191.
- Eifert, G. H. (in press). The acquisition and cognitive-behavioral therapy of phobic anxiety. In I. M. Evans & G. H. Eifert (Eds.), *Unifying behavior therapy: Contributions of paradigmatic behaviorism* (pp. 169–183). New York: Springer.
- Eifert, G. H., & Lauterbach, W. (in press). Relationships between overt behavior to a fear stimulus and self-verbalizations measured by different assessment strategies. *Cognitive Therapy and Research*.
- Eifert, G. H., & Schermelleh, K. (1985). Language conditioning, emotional instructions, and cognitions in conditioned responses to fear-relevant and fear-irrelevant stimuli. *Journal of Behavior Therapy and Experimental Psychiatry*, *16*, 101–110.
- Ellis, A. (1979). A note on the treatment of agoraphobics with cognitive mediation versus prolonged exposure in vivo. *Behaviour Research and Therapy*, *17*, 162–164.
- Ellis, A. (1983). The philosophic implications and dangers of some popular behavior therapy techniques. In M. Rosenbaum, C. M. Franks, & Y. Jaffe (Eds.), *Perspectives on behavior therapy in the Eighties*. New York: Springer.
- Eysenck, H. J. (1960). *The structure of human personality*. New York: Macmillan.
- Eysenck, H. J. (1964). *Crime and personality*. London: Routledge & Kegan Paul.
- Eysenck, H. J. (1982). Neobehaviouristic (S-R) theory. In G. T. Wilson & C. M. Franks (Eds.), *Contemporary behavior therapy: Conceptual and empirical foundations*. (pp. 205–276). New York: Guilford Press.
- Evans, I. M. & Weiss, A. R. (1978). Process studies in language conditioning II: The role of semantic relevance in conditioning negative emotional responses. *Journal of Behavior Therapy and Experimental Psychiatry*, *9*, 121–124.
- Gupta, B. S. (1973). The effects of stimulant and depressant drugs on verbal conditioning. *British Journal of Psychology*, *64*, 553–557.
- Hafner, J. L., & Linkenhoker, D. (1974). Quantifying the reinforcement value of verbal items on drug addicts that were inmates of a prison. *Journal of the Association for the Study of Perception*, *9*, 17–25.
- Hekmat, H. (1972). The role of imagination in semantic desensitization. *Behavior Therapy*, *3*, 223–231.
- Hekmat, H. (1974). Three techniques for reinforcement modification: A comparison. *Behavior Therapy*, *5*, 541–548.
- Hekmat, H. (1977). Semantic behavior therapy: Unidimensional or multidimensional. *Behavior Therapy*, *8*, 805–809.
- Hekmat, H., & Lee, Y. B. (1970). Conditioning of affective self-references as a function of semantic meaning of verbal reinforcers. *Journal of Abnormal Psychology*, *76*, 427–433.
- Hekmat, H., & Theiss, M. (1971). Self-actualization and modification of affective self-disclosures during a social conditioning interview. *Journal of Counseling Psychology*, *18*, 101–105.
- Hekmat, H., & Vanian, D. (1971). Behavior modification through covert semantic desensitization. *Journal of Consulting and Clinical Psychology*, *36*, 248–251.
- Hekmat, H., Deal, R., & Lubitz, R. (1985). Instructional desensitization: A semantic behavior treatment of anxiety disorder. *Psychotherapy: Theory, Research, Practice, Training*, *22*, 273–280.
- Highlen, P. S., & Nicholas, R. P. (1978). Effects of locus of control, instructions, and verbal conditioning on self-referenced affect in a counseling interview. *Journal of Counseling Psychology*, *25*, 177–183.
- Izard, C., Kagan, J., & Zajonc, R. (Eds.). (1984). *Emotion, cognition, and behavior*. Cambridge: Cambridge University Press.
- Javierto, S. (1971). Verbal conditioning, awareness, and intelligence. *Philippine Journal of Psychology*, *4*, 88–96.
- Lilliston, L. (1972). Verbal conditioning of value-related material. *Journal of Consulting and Clinical Psychology*, *39*, 478–482.
- Luria, A. (1961). *The role of speech in the regulation of normal and abnormal behaviors*. New York: Liveright.
- Mahoney, M. J. (1977). Reflections on the cognitive learning trend in psychotherapy. *American Psychologist*, *32*, 5–13.
- Mahoney, M. J. (1979). Cognitive and non-cognitive views in behavior modification. In P. D. Sjoden, S. Bates, & W. S. Dockens (Eds.), *Trends in behavior therapy*. New York: Academic Press.
- Maltzman, I. (1977). Orienting in classical conditioning and generalization of the galvanic skin response to words: An overview. *Journal of Experimental Psychology: General*, *106*, 111–119.

- Marshall, W. L. (1985). The effects of variable exposure in flooding therapy. *Behavior Therapy, 16*, 117–135.
- Martin, D. G. (1975). *Learning-based client-centered therapy*. Monterey: Brooks/Cole.
- Martin, M., & Levey, A. B. (1978). Evaluative Conditioning. *Advances in Behaviour Research and Therapy, 1*, 57–102.
- Martin, I. M., & Levey, A. B. (1985). Conditioning, evaluations and cognitions: An axis for integration. *Behaviour Research and Therapy, 23*, 267–275.
- Master, S., & Gershman, L. (1983). Physiological responses to rational-emotive self-verbalizations. *Journal of Behavior Therapy and Experimental Psychiatry, 14*, 289–296.
- Meichenbaum, D. (1969). The effects of instructions and reinforcement on thinking and language behavior of schizophrenics. *Behaviour Research and Therapy, 7*, 101–114.
- Meichenbaum, D. (1977). *Cognitive-behaviour modification: An integrative approach*. New York: Plenum Press.
- Meichenbaum, D., & Cameron, R. (1973). Training schizophrenics to talk to themselves: A means of developing attentional controls. *Behavior Therapy, 4*, 515–534.
- Meichenbaum, D., & Cameron, R. (1982). Cognitive behavior therapy. In G. T. Wilson & C. M. Franks (Eds.), *Contemporary behavior therapy: Conceptual and empirical foundations*. New York: Guilford Press.
- Miller, P. M., & Drennen, W. T. (1970). Establishment of social reinforcement as an effective modifier of verbal behavior in chronic psychiatric patients. *Journal of Abnormal Psychology, 76*, 392–395.
- Mohan, V., & Dharmani, I. (1979). The effect of intelligence and personality on verbal conditioning. *Psychologica Belgica, 16*, 223–232.
- Namének, A. A., & Schuldt, W. J. (1971). Differential effects of experimenters' personality and instructional sets on verbal conditioning. *Journal of Counseling Psychology, 18*, 170–172.
- Nydegger, R. V. (1972). The elimination of hallucinatory and delusional behavior by verbal conditioning and assertive training: A case study. *Journal of Behavior Therapy and Experimental Psychiatry, 3*, 225–227.
- Öhman, A. (1979). Fear relevance, autonomic conditioning, and phobias: A laboratory model. In P. D. Sjoden, S. Bates, & W. S. Dockens (Eds.), *Trends in behavior therapy*. New York: Academic Press.
- O'Leary, K. D., & Wilson, G. T. (1980). *Principles of behavior therapy*. Englewood-Cliffs, NJ: Prentice-Hall.
- Pansa, M. (1979). Verbal conditioning of affect responses of process and reactive schizophrenics in a clinical interview situation. *Journal of Medical Psychology, 52*, 175–182.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.
- Rachman, S. J. (1981). The primacy of affect: Some theoretical implications. *Behaviour Research and Therapy, 19*, 279–290.
- Razran, G. (1939). A quantitative study of meaning by a conditioned salivary technique. *Science, 90*, 89–90.
- Razran, G. (1965). Russian physiologists' psychology and American experimental psychology. *Psychological Bulletin, 62*, 42–64.
- Razran, G. (1971). *Mind in evolution*. Boston; MA: Houghton Mifflin.
- Rhodes, R., Shames, G., & Egolf, D. (1971). Awareness in verbal conditioning of language themes during therapy with stutterers. *Journal of Communication Disorders, 4*, 30–39.
- Rogers, T., & Craighead, W. E. (1977). Physiological responses to self-statements. *Cognitive Therapy and Research, 1*, 99–120.
- Rosenbaum, M. (1983). Learned resourcefulness as a behavioral repertoire for the self-regulation of internal events. In M. Rosenbaum, C. M. Franks, & Y. Jaffe (Eds.), *Perspectives on behavior therapy in the eighties*. New York: Springer.
- Ross, A. O. (1985). To form a more perfect union: It is time to stop standing still. *Behavior Therapy, 16*, 195–204.
- Sappington, A. A. (1975). Semantic desensitization: Language conditioning or demand characteristics manipulation? *Perceptual and Motor Skills, 41*, 747–753.
- Sappington, A. A., Burleson, R., Studstill, L., Rice, J., Gordon, J., & Cornelison, K. (1982). Reduction of avoidant behavior through the semantic conditioning of a self-related target stimulus. *Cognitive Therapy and Research, 6*, 315–323.
- Staats, A. W. (1963). *Complex human behavior*. New York: Holt, Rinehart & Winston. (With contributions by C. K. Staats)
- Staats, A. W. (1968). *Learning, language and cognition*. New York: Holt, Rinehart & Winston.

- Staats, A. W. (1972). Language behavior therapy: A derivative of social behaviorism. *Behavior Therapy, 3*, 165–192.
- Staats, A. W. (1975). *Social behaviorism*. London: Irvin-Dorsey Press.
- Staats, A. W. (1980). Behavioral interaction and interactional psychology theories of personality: Similarities, differences, and the need for unification. *British Journal of Psychology, 71*, 205–220.
- Staats, A. W., & Heiby, E. M. (1985). Paradigmatic behaviorism's theory of depression: Unified, explanatory, and heuristic. In S. Reiss & R. R. Bootzin (Eds.), *Theoretical issues in behavior therapy*. New York: Academic Press.
- Truax, C. B. (1966). Reinforcement and nonreinforcement in Rogerian psychotherapy. *Journal of Abnormal and Social Psychology, 71*, 1–9.
- Tryon, W. W., & Briones, R. G. (1985). Higher-order semantic counterconditioning of Filipino women's evaluations of heterosexual behaviors. *Journal of Behavior Therapy and Experimental Psychiatry, 16*, 125–131.
- Weiss, A. R., & Evans, I. M. (1978). Process studies in language conditioning—I: Counterconditioning of anxiety by “calm” words. *Journal of Behavior Therapy and Experimental Psychiatry, 9*, 115–119.
- Wilson, G. T., & Evans, I. M. (1977). The therapist-client relationship in behavior therapy. In A. S. Gurman & A. M. Razin (Eds.), *The therapist's contribution to effective psychotherapy*. New York: Pergamon Press.

PART III

COGNITIVE THEORY

CHAPTER 9

Trait Anxiety and Cognition

Michael W. Eysenck and Andrew Mathews

INTRODUCTION

This chapter should be read in conjunction with the following chapter by Mathews and Eysenck. The emphasis differs in the two chapters, in that this chapter is concerned with trait anxiety in normals, whereas the chapter by Mathews and Eysenck deals with clinical anxiety. Despite this difference, there is much overlap between the two chapters. As will be discussed in the following, trait anxiety may well predispose to clinical anxiety. In addition, the cognitive differences between normals high and low in trait anxiety resemble those between patients with generalized anxiety and normal controls.

It seems likely that there are pronounced individual differences in terms of susceptibility to anxiety neurosis, and numerous dimensions of individual differences are probably involved. However, the single most obvious one is trait anxiety, which was defined by Spielberger, Gorsuch, and Lushene (1970) as "relatively stable individual differences in anxiety proneness" (p. 3). There are several self-report questionnaire measures of trait anxiety available, such as Spielberger's State Trait Anxiety Inventory and Taylor's Manifest Anxiety Scale. In addition, neuroticism as measured by the Eysenck Personality Questionnaire and Byrne's Repression-Sensitization Scale correlate highly with standard tests of trait anxiety, and so can be regarded as alternative measures.

It is common in the literature to distinguish between trait anxiety and state anxiety. Whereas the former is a semipermanent personality dimension, the latter is the usually rather transient experience of feeling anxious. According to Spielberger *et al.* (1970), state anxiety is "characterized by subjective, consciously perceived feelings of tension and apprehension, and heightened autonomic nervous system activity" (p. 3). The distinction between trait and state anxiety is conceptually clear and of relevance to the research to be discussed. However, there are some difficulties

MICHAEL W. EYSENCK • Department of Psychology, Birkbeck College, University of London, London WC1E 7HX, England. ANDREW MATHEWS • Department of Psychology, St. George's Hospital Medical School, Cranmer Terrace, London SW 17, England.

at an empirical level, because measures of trait and state anxiety typically correlate quite highly with each other. In addition, anxious patients may remain anxious for a period of a few months; such anxiety is neither relatively transient like state anxiety nor semipermanent like trait anxiety. Despite these limitations, the distinction between trait and state anxiety has proved useful in a number of contexts.

Trait anxiety is determined in part by genetic factors. For example, consider an unusually thorough study carried out on 12,898 twin pairs by Floderus-Myrhed, Pedersen, and Rasmusson (1980). The estimated heritability for neuroticism was .50 and .58 for males and females, respectively. Findings like these led Eysenck and Eysenck (1985) to conclude as follows: "No serious worker in this field denies that genetic factors account for at least something like half of the variance" (p. 96). In line with the hereditarian position is the reasonable longitudinal consistency shown by trait anxiety or neuroticism. When account is taken of the relatively low period-free reliability of most questionnaire measures of trait anxiety, then their mean annual stability is approximately .98 (Conley, 1984).

If it is true that trait anxiety predisposes to anxiety neurosis, then individuals who have suffered from anxiety neurosis should on average have had higher pre-morbid levels of trait anxiety than normal controls. The best test of this hypothesis would be a large-scale prospective study, in which trait anxiety was assessed on a normal sample and those who subsequently suffered from anxiety neurosis were compared with those who did not. Because such a study has never been carried out, we must rely on indirect evidence. For example, there is a study of obsessive-compulsive neurosis by McKeon, Roa, and Mann (1984) that is discussed more fully in the following chapter. Their findings suggested that those who are high in trait anxiety are in some sense more vulnerable and susceptible to stress.

An alternative approach is to assess trait anxiety in individuals who have recovered from anxiety neurosis. The value of this approach depends on the extent to which trait anxiety levels after recovery correspond to pre-morbid levels. It is possible, for example, that the experience of being clinically anxious means that trait anxiety is more or less permanently elevated above its pre-morbid levels, but there is no real evidence for this. Ingham (1966) administered the Maudsley Personality Inventory to a sample of 119 neurotic patients, and then readministered the test 3 years later. He found that the extent of the reduction in neuroticism scores at the second testing correlated with the degree of recovery shown. Those recovered patients who were "very much better" had neuroticism scores that resembled those of a normal random sample. However, rather different findings were obtained from those former patients whose improvement was between "slightly better" and "very much better" as assessed by a relative. The mean neuroticism score among these former patients was 37.3 for males (against a mean for normals of 25.4), and it was 33.9 for females (against 29.3 for normals). These findings are at least consistent with the hypothesis that high trait anxiety or neuroticism predisposes to anxiety neurosis.

Any serious attempt to provide an adequate theoretical understanding of trait anxiety should start from the premise that a number of partially separate but interdependent systems are involved. For example, Lang (1971) suggested a tripartite division into behavioral, physiological, and verbal data. The notion that these different kinds of data reflect the functioning of somewhat separate systems is supported

by the weight of evidence indicating that they often fail to respond concordantly (e.g., Craske & Craig, 1984; Weinberger, Schwartz, & Davidson, 1979).

Some of these theoretical distinctions among anxiety systems have been incorporated into questionnaire measures. The Cognitive-Somatic Anxiety Questionnaire (Schwartz, Davidson, & Goleman, 1978), as the name implies, provides separate measures of cognitive and somatic anxiety. In similar fashion, Liebert and Morris (1967) extracted separate measures of worry and emotionality from a test-anxiety scale.

It is unfortunate that many theories of trait anxiety or neuroticism have been concerned almost exclusively with a single anxiety system. Thus, for example, Hamilton (1983) proposed a theory of trait anxiety that was almost entirely cognitive. He virtually identified trait anxiety with information stored in long-term memory. In contrast, Gray (1982) placed great emphasis on the physiological system in his theory of anxiety. According to him, anxiety consists of activity in the behavioral inhibition system, which comprises the septo-hippocampal system, its monoaminergic afferents from the brain stem, and its neocortical projection in the frontal lobe.

What is required is of great complexity. Firstly, the particular contributions of each system to anxiety need to be investigated thoroughly. Secondly, the interrelationships among the cognitive, physiological, and behavioral systems must be considered. There has been very little research on these interrelationships as yet. Most of the research discussed in this chapter was based on the implicit assumption that the cognitive system can be decoupled from the other systems involved in anxiety. This may be a convenient fiction, but it is important to remember that it is a fiction.

THE COGNITIVE SYSTEM: OVERVIEW

It has often not been recognized that research on cognitive differences between those high and low in trait anxiety has followed a number of approaches. The approach that has been used most extensively involves comparing the performance of individuals high and low on trait anxiety on a variety of cognitively demanding tasks. The basic data obtained indicate whether anxiety has improved or impaired task performance, and these data are then used to infer the internal processes that are alleged to be responsible.

In spite of the popularity of this approach, it suffers from significant limitations. Some of these limitations relate to the oversimplified theoretical formulations that have been proposed, and are discussed at some length in the next section. At a more general level, research based on this approach may prove to have relatively modest implications for therapy. In part, this is because issues that are clinically relevant, such as the nature of the stimuli that trigger anxiety and the coping strategies used to handle stress, are largely ignored.

The approach we have just discussed is concerned with the effects of anxiety on current internal processes. A very different approach focuses on possible differences in the content and organization of long-term memory as a function of trait anxiety: in other words, the emphasis is on structural as well as processing differences. Because

the subjective experiences of individuals high and low in trait anxiety are very different, there should be predictable consequences for the information stored in long-term memory. Differences in the content of long-term memory as a function of trait anxiety would be expected to influence many psychological processes, including perception, attention, learning, and memory.

A third approach is based on the assumption that the content of task stimuli is an important factor when considering cognitive differences between those high and low in trait anxiety. The basic assumption underlying this approach is that high-anxiety and low-anxiety subjects are likely to differ considerably more in their processing of threatening or threat-related stimuli than in their processing of nonthreatening or neutral stimuli. This approach has advantages from the clinical perspective in that it focuses on how the kinds of stimuli that produce anxiety are initially processed, and also considers individual differences in subsequent coping with such stimuli.

In sum, as Eysenck (in press-a) pointed out, cognitive differences as a function of trait anxiety have been investigated in a number of different ways. Broadly speaking, the three major research approaches have focused on processing differences, structural differences, and the effects of stimulus content, respectively. All of these approaches are discussed more fully in the sections that follow. An important issue concerns the links that exist among these three approaches. Thus, for example, it is entirely reasonable to suggest that differences in long-term memory between those high and low in trait anxiety may affect processing efficiency and the reactions to threatening stimuli. Conversely, inefficient processing and inadequate reactions to threat by high-anxiety individuals may lead to the storage of information in long-term memory that differs considerably from that of low-anxiety individuals. However, research has so far largely failed to confirm or to deny the existence of such links.

THE COGNITIVE SYSTEM: PERFORMANCE DEFICITS

One of the most obvious effects of anxiety on the cognitive system is to disrupt the thought processes involved in dealing with the problems of everyday life. Comparisons between individuals high and low in trait or test anxiety under laboratory conditions have frequently revealed performance deficits associated with high anxiety. The fact that cognitive processes and task performance are often disrupted by anxiety may have implications for the genesis of anxiety neurosis. The poor performance caused by anxiety may well give rise to feelings of failure and of an inability to cope, which then lead to increased anxiety. Thus, a vicious circle may develop, in which anxiety impairs performance, and impaired performance increases anxiety and tension.

The literature on anxiety and cognitive task performance has been reviewed recently by Eysenck (1984a) and by Eysenck and Eysenck (1985). It indicates very clearly that the effects of trait anxiety vary from task to task, and that it is mainly relatively difficult or demanding tasks that are adversely affected by high trait anxiety. For example, Mayer (1977) found that individuals high in trait anxiety were much less successful than those low in trait anxiety in solving complex cognitive problems

(e.g., water-jar problems; anagrams), but there were no effects of trait anxiety on simple tasks such as visual search. Therefore, an important theoretical issue is to identify the processes and mechanisms that produce interactions between trait anxiety and the nature of the task.

The adverse effects of trait anxiety on performance are often mediated by state anxiety. Thus, for example, low-anxiety subjects are more likely to outperform high-anxiety subjects in stressful (e.g., ego-involving instructions; failure feedback) conditions than in nonstressful conditions. In recent research, Leon and Revelle (1985) discovered that performance on an analogical reasoning task was significantly affected by state anxiety but not by trait anxiety.

Easterbrook (1959) made an impressive attempt to account for the interaction between anxiety level and task difficulty in attentional terms. He argued that states of high anxiety, emotionality, and arousal all produce a reduction in the range of cue utilization that can be regarded as a narrowing of attention. If difficult tasks involve more cues than easy ones, then high-anxiety individuals are at a disadvantage on difficult tasks because they focus on only some of the many task cues that need to be considered. This hypothesis is inadequate because it assumes that high anxiety leads to great concentration on some of the task cues or stimuli. The implicit assumption that anxiety reduces distractibility conflicts with clinical observation; as Korchin (1964) pointed out, "The anxious patient is unable to concentrate, hyper-responsive, and hyper-distractible." At an experimental level, Dornic and Fernaus (1981) reported that the task performance of neurotic introverts (i.e., individuals high in trait anxiety) was more adversely affected than that of stable extraverts (i.e., individuals low in trait anxiety) by distracting stimuli in each of three experiments.

It may be possible to incorporate parts of Easterbrook's (1959) hypothesis into a more viable theoretical formulation. According to Wachtel (1967), attention can be compared to a beam of light, and it is important to distinguish between the width of the beam and movements of the beam. Easterbrook (1959) may have been correct in his assumption that anxiety narrows the attentional beam, but mistaken in believing that there are few movements of the beam in states of anxiety. It is more likely that anxiety causes the beam to roam widely throughout the perceptual field, as is suggested by the anxious individual's distractibility and poor concentration.

The basic strategy of comparing the performance of individuals high and low in trait anxiety across a range of tasks is a limited one. The data that are obtained are usually interpreted by means of a simplistic one-stage model based on the assumption that task performance directly reflects internal processes. Thus, if trait anxiety does not affect task performance, it is concluded that trait anxiety has had no effect on internal processes. The clearest refutation of this line of argument comes in a study by Weinberg and Hunt (1976). They discovered that trait anxiety had no effect on performance of a throwing task prior to the introduction of feedback. However, they also used electromyography to provide various measures of muscle activity during the throwing task, and these data pointed to a very different conclusion:

High-anxious subjects anticipated significantly longer with the agonists and shorter with the antagonists than did the low-anxious group. Therefore, they were preparing for the throw in all of the muscles while low-anxious subjects were preparing mostly with the

antagonist muscles. This implies that high-anxious subjects were using more energy than necessary, and expending it over a greater period of time, than were low-anxious subjects. (p. 233)

An alternative theoretical approach has been proposed by Eysenck (1979, 1982, 1984a, in press-a). He argued that the traditional one-stage model should be replaced by a two-stage model, in which task performance reflects not only the natural effects of anxiety on internal processes, but also the more or less successful attempts to compensate for any adverse effects of anxiety. It is important within this conceptualization to distinguish between performance effectiveness (i.e., the quality of performance) and processing efficiency, which refers to the relationship between the effectiveness of performance and the amount of effort or processing resources invested in it. The available evidence indicates that individuals high in trait anxiety often have lower processing efficiency than individuals low in trait anxiety, but trait anxiety less frequently affects performance effectiveness (see Eysenck, 1982, for a review). Thus, as in the study by Weinberg and Hunt (1976), nonsignificant effects of trait anxiety on performance effectiveness often camouflage genuine effects on processing efficiency.

Recent empirical evidence has provided support for the notion that individuals high in trait anxiety endeavor to compensate for the detrimental effects of anxiety by increased effort or investment of processing resources under normal circumstances. Because those high in trait anxiety are closer to maximum resource allocation than those low in anxiety, they should benefit less from a manipulation designed to increase effort. Eysenck (1985) and Calvo (1985) found that monetary incentives improved the performance of low-anxious subjects but had no effect on the performance of high-anxious subjects.

Why do individuals high in trait anxiety tend to have reduced processing efficiency? According to Eysenck (1979, 1982), their high level of state anxiety disrupts the functioning of working memory (cf. Baddeley & Hitch, 1974), a system that is concerned with the processing of task information and the storage for short periods of time of task-relevant information. The available evidence (reviewed by Eysenck, 1982) suggests that state anxiety mainly affects the central executive component of working memory. The central executive is modality free, has limited capacity, and resembles attention. The typical interaction between anxiety and task difficulty can be accounted for by this theory, if we assume that the critical dimension of task difficulty is the demands placed on the central executive component of the working memory system.

There is little direct experimental evidence relevant to the issue of why it is that high state anxiety impairs the functioning of the central executive. However, there have been a number of studies in which separate measures of the worry and emotionality components of anxiety have been obtained and related to task performance. The typical finding is that poor task performance is associated more with worry than with emotionality (Morris, Davis, & Hutchings, 1981). This suggests that task-irrelevant processing relating to concerns about performance and negative self-evaluations may be preempting some of the resources of the central executive. In other words, information about personal inadequacies may be accorded priority of processing over task information by individuals high in trait anxiety. However, all

that has been established is that worry and performance are correlated, and so the direction of causality remains unclear. It is possible that worry is an effect of poor performance (or poor anticipated performance) rather than a cause of it (Klinger, 1985), but it seems more likely that there are actually bidirectional effects of worry on performance, and of performance on worry.

THE COGNITIVE SYSTEM: LONG-TERM MEMORY

Whereas there may be several differences in the cognitive systems of those high and low in anxiety, especially important differences are likely to relate to the contents and organization of long-term memory. Hamilton (1983) subscribed to this hypothesis in a typically trenchant fashion:

Anxiety should be regarded as a particular set or network of connotative data that, on the basis of past experience and autonomous elaboration of their cognitive structures, provides a store of long-term memories . . . The greater the predisposition to generate aversive expectancies or behavior outcomes, the greater the appropriate memory store, the lower the retrieval threshold for this type of information, and the greater the response bias toward primary processes of identifying and avoiding real or potential aversiveness. (p. 114)

In general terms, this emphasis on long-term memory seems appropriate. Long-term memory may vary as a function of trait anxiety because of the very different subjective experiences of high- and low-anxiety individuals. For example, Williams (1981) measured mood on 12 days in a 26-day period, and discovered that neuroticism correlated $-.73$ with average mood. The relevance of this to long-term memory becomes clear if we consider the phenomenon of mood-congruent learning (Bower, Gilligan, & Monteiro, 1981). The essence of this phenomenon is that emotionally toned information is learned best when there is correspondence between its affective value and the learner's current mood. Because someone high in trait anxiety is usually in a more negative mood than is someone low in trait anxiety, the implication of the mood-congruent learning effect is that his or her long-term memory will contain much more negatively toned information.

Individual differences in mood state are also relevant to the retrieval of information from long-term memory. The phenomenon of mood-state-dependent retrieval (Bower, 1981) means that information acquired in one mood state is more likely to be recalled at a later time if the mood state at recall is similar to the mood state at acquisition rather than different. This phenomenon has sometimes proved difficult to replicate (e.g., Bower & Mayer, 1985), but recent evidence indicates that it can be obtained consistently provided that the learner perceives a causal relationship between the information that must be learned and his or her current mood (Bower, 1985).

Somewhat more direct evidence that long-term memory may differ between those high and low in trait anxiety was obtained by Mayo (1983). He asked subjects to retrieve one specific real-life personal experience from memory to each of a series of stimulus words (e.g., children, smell, swimming). Neuroticism correlated $+.39$ with the number of unpleasant memories retrieved, and the size of the correlation was virtually unaffected when mood state at the time of testing was factored out.

Although these findings may reflect individual differences in retrieval strategy, it is at least as plausible to argue that they are due to genuine differences in the nature of the information stored.

Other studies have considered long-term memory for negatively toned information acquired in the laboratory. Young and Martin (1981) presented positive and negative trait words that subjects were led to believe described their own personality. Those high in neuroticism showed a selective negative bias in recall, recalling negative rather than positive information to a greater extent than those low in neuroticism. This finding was replicated by Martin, Ward, and Clark (1983), who also discovered that there was no selective negative bias in recall for trait words descriptive of "a typical undergraduate from your college." They argued that the selective negative bias in recall of personally relevant trait terms occurred because those high in neuroticism selectively attended to negative information about themselves at the time of acquisition. Thus, this selective negative bias may resemble the phenomenon of mood-congruent learning (Bower *et al.*, 1981). However, matters are complicated by the fact that it has not proved possible to obtain a selective negative bias in recall with generalized anxiety patients (Mogg, Mathews, & Weinman, in press), research which is discussed in greater detail in the next chapter.

There are various possible reasons for this apparent discrepancy in the data. One is that neuroticism, which is an approximate measure of trait anxiety, also correlates with measures of depression. It is thus possible that the selective negative bias in recall obtained by Young and Martin (1981) and by Martin *et al.* (1983) is attributable to depression rather than trait anxiety *per se*, although Martin *et al.* (1983) reject this interpretation of their data.

When discussing the effects of trait anxiety on the contents of long-term memory, it is important to consider worries. Borkovec, Robinson, Pruzinsky, and DePree (1983) reported a correlation coefficient of $+ .67$ between trait anxiety and the percentage of a typical day that was spent worrying. This association can be interpreted in a number of different ways. It is possible that the number and structure of worries in long-term memory do not vary as a function of trait anxiety, and that these differences in the amount of worrying are simply a reflection of mood-state-dependent retrieval effects. On the other hand, it may be the case that those high in trait anxiety have a greater number of well-organized worries than those who are low in trait anxiety.

Some relevant evidence was obtained by Eysenck (1984b). He obtained the usual greater incidence of worrying among high-anxiety individuals even when the high- and low-anxiety groups had comparable initial levels of state anxiety. This is consistent with a structural rather than a process interpretation of the association between anxiety and worry.

At present, relatively little is known about worry and worrying. The first essential is probably to understand more about the underlying structure of worries. A start in this direction was made by Mathews and MacLeod (unpublished raw data). They prepared a 100-item questionnaire incorporating most common worries and concerns, and asked their subjects to indicate the frequency with which they experienced each worry. Cluster analysis revealed the existence of eight major categories of worry. These are shown in Table 1, along with representative items.

TABLE 1. The Major Categories of Worry Revealed by a Cluster Analysis of a 100-Item Worry Questionnaire"

Category	Representative items
1. Coping with responsibility	Forced to take on too much responsibility
2. Relationships	Partner/spouse may not love you
3. Social insecurity and rejection	More stupid than you realize
4. Lack of personal fulfilment	Are an insignificant cog in society which nobody cares about
5. Financial troubles	Could become financially poorer in future
6. Societal and environmental problems	Perhaps pollution is more serious than we are told
7. Personal danger	You could get burgled
8. Panic/loss of control	Might freeze, through panic, in a public place

"From Mathews and MacLeod (unpublished raw data).

Although the worry questionnaire was given to anxious patients and to normals, the numbers were insufficient to compare the worry structures of the two groups. However, it is interesting to note that anxious patients scored slightly *lower* than normals on Societal and Environmental Problems, although they scored much higher in every other worry category.

Individuals high and low in trait anxiety might differ from each other in their worries in various ways. These include at least the number of worry categories that can be identified, the nature of those worry categories, and the organization of worries. Direct evidence on these issues is lacking, but the procedures adopted by Mathews and MacLeod (unpublished raw data) could certainly provide relevant evidence. However, the methodological problems with self-report data point to the use of converging operations, in which worries are investigated by a combination of self-report and experimental techniques (cf. Coltheart & Evans, 1981).

One way in which long-term memory can be conceptualized is as a hierarchically structured system with very general and rather abstract concepts at the top of each hierarchy and rather specific, concrete concepts at the bottom. Within this conceptualization, many worries can be regarded as relatively specific units of information, and can be contrasted with broad memory structures such as general schemata. An example of the kind of schematic differences that may exist between high- and low-anxiety groups was proposed by Butler and Mathews (1983), in a study discussed further in the following chapter. They argued that anxious patients (but not normal controls) possess "danger schemata." This argument was substantiated by the fact that the patients reported that they were significantly more at risk from potential environmental dangers than other people, whereas the normal controls did not produce this self-other discrepancy.

It appears probable that there are close links between trait anxiety and the structures and processes of long-term memory, including general schemata. There are various important implications. It has been established (e.g., Bransford, 1979) that the processing and comprehension of presented stimuli are much affected by the prior knowledge and experience available in long-term memory. In particular, there are significant differences between the stimulus as presented and the stimulus

as encoded because of the inferential processes that have developed as a result of prior knowledge. At the time of retrieval, as Bartlett (1932) pointed out, schemata stored in long-term memory may influence (and even distort) the nature of the information that is recalled.

THE COGNITIVE SYSTEM: REACTIONS TO THREAT

Various differences in cognitive functioning between those high and low in trait anxiety have already been discussed. So far, however, we have not considered in detail the clinically relevant issue of how individuals differ in their processing of threatening stimuli. An intriguing theoretical formulation was proposed by Byrne (1964). He argued quite simply that there are two main possible reactions to threatening stimuli: they may either be approached (i.e., thoroughly processed) or avoided (i.e., minimally processed). He then claimed that some individuals (whom he referred to as sensitizers) exhibit a consistent tendency to approach threatening stimuli, whereas other individuals (repressors) typically avoid such stimuli. The Repression-Sensitization Scale (Byrne, 1961) was specifically designed to test these individual differences.

Byrne (1964) did not emphasize the point, but it seems fairly obvious that his hypotheses are relevant only to situations in which mild threats are presented. It is most unlikely that anyone would adopt an avoidance strategy if confronted by a major threat to life or health. Under such circumstances, nearly everyone would allocate all of their available resources to the source of the threat.

The relevance of Byrne's (1964) theoretical distinction between repressors and sensitizers to trait anxiety needs clarification at this point. Byrne (1964) himself argued that the Repression-Sensitization Scale was not simply a measure of trait anxiety, but the empirical evidence does not support that contention. As was mentioned earlier, scores on the Repression-Sensitization Scale correlate so highly with those on tests of trait anxiety (Watson & Clark, 1984), that one must conclude that they are equivalent measures of the same construct. Therefore, repressors are low-anxiety individuals and sensitizers are high-anxiety individuals.

If Byrne (1964) is correct in his assertion that sensitizers or those high in trait anxiety engage in excessive processing of threatening stimuli, then clearly the environment as subjectively experienced by them is more threatening and dangerous than it is for repressors or those low in trait anxiety. Thus, these processing strategies should help to explain why high trait anxiety predisposes to anxiety neurosis. However, before jumping to any such conclusions, it is obviously necessary to evaluate the relevant experimental evidence.

Much of the published literature on the processing of threatening stimuli by repressors and sensitizers has been reviewed by Byrne (1964) and by Krohne and Rogner (1982). In essence, a rather inchoate picture emerges, and numerous findings are inconsistent with the predictions of Byrne's (1964) hypothesis. For example, several studies have investigated the phenomenon of perceptual defense, in which emotionally threatening or taboo stimuli have higher perceptual recognition thresholds than neutral stimuli (Dixon, 1981). If perceptual defense occurs because there is reduced processing, or avoidance, of threatening stimuli, then it follows that there

should be stronger perceptual defense effects among repressors than among sensitizers. In fact, many studies have failed to support this prediction (for reviews, see Eysenck, *in press-a, b*). Moreover, it is not clear that those studies reporting a greater perceptual defense effect for repressors than for sensitizers can be interpreted unequivocally as supporting Byrne (1964). The reason is that no attempt was made to decide whether the difference between repressors and sensitizers in the size of the perceptual defense effect was due to genuine perceptual factors or to individual differences in response bias. It is interesting to note that in the two studies in which perceptual sensitivity uncontaminated by response bias was assessed (Van Egeren, 1968; Wagstaff, 1974), there were no discernible differences between repressors and sensitizers in perceptual defense.

Why do the predicted differences between repressors and sensitizers in perceptual defense fail to materialize in the perceptual defense paradigm? One obvious answer is simply that Byrne's (1964) hypothesis is incorrect, and that there are, in fact, no consistent differences between low-anxiety and high-anxiety individuals in their initial processing of threatening stimuli. However, it is also possible that Byrne's (1964) hypothesis applies only to certain situations, and that these situations do not include the perceptual defense paradigm. For example, perceptual recognition may require primarily early or low-level perceptual processes, whereas the approach and avoidance strategies identified by Byrne (1964) may operate only on later or high-level perceptual processes. However, some evidence is difficult to reconcile with this view. Carroll (1972) and Lewinsohn, Berquist, and Brelje (1972) presented threatening and neutral visual stimuli for periods of time that gave ample opportunity for the use of high-level perceptual processes. Despite this, measures of the duration of visual attention to the threatening stimuli (e.g., pictures of mutilated bodies and corpses) failed to reveal any differences between repressors and sensitizers.

There is a further hypothesis that deserves serious attention. It is possible that systematic approach and avoidance tendencies are manifest only when processing resources must be allocated to two or more concurrent stimuli. Because virtually all of the perceptual defense and duration of visual attention studies have involved the presentation of only one stimulus at a time, it follows that this selective allocation of resources could not occur in those studies.

In order to test this selective bias interpretation, it is necessary to measure the allocation of processing resources in a situation where at least one threatening and one neutral stimulus are presented concurrently. The first such study was carried out by a student of the senior author (C. Halkiopoulos), using a modified dichotic listening task in which pairs of words were presented concurrently, one to each ear. All of the words presented to one ear had to be shadowed, that is, reported back aloud. Threatening (e.g., grave, fail) and nonthreatening words were presented on the shadowed or attended ear, whereas only neutral words were presented to the unattended ear. The allocation of processing resources was assessed by requiring the subjects to respond as quickly as possible to tones that were occasionally presented to the shadowed or unattended ear shortly after a pair of words had been presented.

The Facilitation-Inhibition Scale (Ullmann, 1962) was administered to the subjects. This scale correlates very highly with the Repression-Sensitization Scale. There were quite large differences between inhibitors and facilitators in their mean

latencies to tones following threatening and neutral words, and overall there was a highly significant interaction involving facilitation-inhibition, attended word type, and probe channel. Exactly as predicted by Byrne (1964), facilitators (i.e., those high in trait anxiety) responded very quickly when the probe followed a threatening word in the same ear, and very slowly when it came after a threatening word in the other ear. Thus, they allocated extra processing resources to the ear on which a threatening word had been presented. Inhibitors showed the opposite pattern of findings, indicating avoidance of the channel on which a threatening word had been presented. Before one attaches too much weight to these findings, it is obviously important that they should be replicated. This has been done by Broadbent (personal communication) and by Eysenck, Halkiopoulos, MacLeod, and Mathews (in preparation) using a visual analog of Halkiopoulos's paradigm with subjects high and low in trait anxiety. Similar findings with clinical patients having a primary diagnosis of generalized anxiety have been reported by MacLeod, Mathews, and Tata (1986), in a study discussed more fully by Mathews and Eysenck (this volume).

There are still various lacunae in our understanding of selective biases in favor or against threatening stimulation, but some issues have been clarified. Firstly, a more complex formulation than the one proposed by Byrne (1964) is required, because consistent individual differences in approach and avoidance strategies are found only in certain circumstances. As a first approximation, these biases are most readily demonstrated when threatening and neutral stimuli are presented concurrently. Thus, these biases are basically selective in nature. Secondly, the tendency of individuals high in trait anxiety to allocate processing resources selectively to threatening stimuli may help to account for their tendency to worry about their performance on demanding tasks (e.g., Eysenck, 1979), and also for their elevated levels of state anxiety even in apparently nonstressful conditions (Watson & Clark, 1984). It is also likely that excessive sensitivity to the minor problems and threats of everyday life might be one of the factors involved in the etiology of anxiety neurosis.

There is some evidence concerning the level of the processing system at which these selective biases operate. Because the relevant research involves clinically anxious patients rather than normals high and low in trait anxiety, it is discussed more fully in the following chapter. The indications are that the selective biases can occur even when there is no conscious awareness of the threatening words, which suggests that these biases are probably operating at a preattentive level.

It remains unclear whether the selective biases are due to trait anxiety *per se*, or whether state anxiety might also play a part. It is also possible, of course, that both trait and state anxiety make separate contributions to selective biases, and data supporting this possibility have been obtained by Broadbent (personal communication).

How do these selective biases develop? At present, any answer to that question must be rather speculative. An attractive possibility is that there is a long process of learning in which initially conscious strategies for coping with threat gradually become automatized. As a consequence, the adult subjects used in the studies previously discussed have developed selective biases of a relatively automatic kind that operate below the level of conscious awareness.

An interesting suggestion (MacLeod, personal communication) is based on the commonsensical view that very mild threats are avoided and greater threats are

approached. Those high in trait anxiety are likely to perceive any given threatening stimulus as more threatening than those low in trait anxiety. They will thus tend to approach and process thoroughly more threatening stimuli than those low in trait anxiety, and thus will develop the approach strategy to a greater extent.

So far we have discussed individual differences in the initial processing of stimuli that are clearly threatening or nonthreatening. The general notion that there are consistent individual differences in approach and avoidance tendencies can also be investigated by considering interpretations of ambiguous stimuli that can be interpreted in either a threatening or a neutral way. The natural expectation is that those high in trait anxiety should select the threatening interpretations of such stimuli more frequently than those low in trait anxiety. If this expectation is supported by the evidence, then there are some possible clinical implications. Someone who consistently perceives ambiguous situations as threatening will obviously regard his or her environment as more threatening than someone who favors neutral interpretations of ambiguity.

There are only a few studies in which interpretation of ambiguity as a function of trait anxiety or repression-sensitization has been investigated. Blaylock (1963) found in one study that sensitizers were more likely than repressors to interpret homographs in a threatening fashion, but this finding was not replicated in a further study. Haney (1973) discovered that sensitizers were much more likely than repressors to interpret ambiguous sentences in threatening ways. Eysenck, MacLeod, and Mathews (in press) simply asked their subjects to write down the spelling of each word presented on a tape recorder. Some of the words were homophones having both a threatening and a neutral meaning (e.g., guilt, gilt). Among normals, the correlation between trait anxiety and the number of threatening interpretations of homophones written down was $+ .60, p < .025$, but state anxiety did not correlate with threatening homophone interpretations.

The threatening interpretations for half of the homophones referred to physical health, and for the other half it referred to social problems. Those individuals who worry mainly about physical health (as assessed by a short questionnaire) might have been expected to differ from worriers about social problems in terms of the kinds of homophones given threatening interpretations, but there was no evidence of this in the data.

The correct explanation of the effects of trait anxiety on homophone interpretation remains unclear. It seems likely that both (or all) meanings of ambiguous stimuli are activated automatically, as is assumed by the exhaustive access model (Simpson, 1984). The meaning that reaches conscious awareness first may then depend on preattentive selective biases that either facilitate or inhibit processing of any threat-related interpretations. Alternatively, it is well established that familiarity plays an important role in the resolution of ambiguity, with more familiar or frequent past interpretations being selected more often than less familiar ones. If, as may well be the case, high-anxiety subjects are more familiar than low-anxiety subjects with the threatening interpretations, then the obtained findings may largely reflect differential familiarity.

For present purposes, it is important to note that both of these theoretical accounts indicate that the cognitive system is involved at an early stage of stimulus

interpretation, and that the functioning of that system varies systematically as a function of trait anxiety. Thus, the cognitive system has a central role to play in deciding whether a stimulus is threatening or not. This decision has implications for the subsequent involvement of the physiological and behavioral systems.

The findings that we have discussed in this section indicate that individuals high in trait anxiety tend to allocate processing resources selectively to threat-related stimuli. They also attach threatening rather than neutral interpretations to ambiguous stimuli. The fact that their cognitive systems function in this way is very much in line with their known susceptibility to stress, and helps to account for their characteristically elevated levels of state anxiety. The findings for low-anxiety individuals are rather more puzzling. It seems natural to argue that those low in trait anxiety are particularly well adjusted on average (Eysenck, 1967). However, the fact that they distinguish between threatening and neutral stimuli at a very early stage of processing, and then systematically avoid the threatening stimuli could be taken to imply that they are much affected by threat, and thus not well adjusted. This position was adopted by Byrne (1964), who argued that individuals who are intermediate on the Repression-Sensitization Scale tend to be less maladjusted than those who are extreme scorers (i.e., repressors or sensitizers).

In fact, low-anxiety individuals seem to constitute a heterogeneous group, including some people who are genuinely free from anxiety and other people who are greatly affected by threat but who are defensive about it. This conclusion is supported by a variety of studies, but the one by Weinberger, Schwartz, and Davidson (1979) is the most convincing. They administered the Marlowe-Crowne Social Desirability Scale, which they claimed was a measure of defensiveness, protection of self-esteem, and inhibition of affect. This questionnaire allowed them to distinguish between repressors (i.e., those with low trait anxiety but high Marlowe-Crowne scores), and low-anxious subjects (i.e., those with low scores on both scales). Three physiological and three behavioral measures were taken during the performance of a stressful phrase-association task, and all six measures indicated that the repressors were significantly more stressed than the low-anxious subjects. The repressors seemed, if anything, to be more stressed than a high-anxiety group, whereas the low-anxious group was the least stressed.

Similar findings were obtained by Gudjonsson (1981). Physiological and self-report measures of stress were taken during a mildly stressful task. Repressors (i.e., those who reported below average self-reported stress but above average physiological stress) tended to have low Marlowe-Crowne scores. Thus, at least some low-anxiety scorers are rather reactive physiologically to stress.

It is a truism in psychology that there are significant individual differences in the perception and interpretation of environmental events. Thus, for example, Koffka (1935) distinguished between the geographical (or objective) environment and the behavioral environment, which "depends upon two sets of conditions, one inherent in the geographical environment, one in the organism" (p. 31). However, it is relatively unusual for individual differences in the behavioral environment to be investigated in a systematic fashion. Byrne (1964) deserves considerable credit for his pioneering theory and research in this area, despite the fact that the approach and avoidance strategies that he identified do not operate over as wide a range of situations

as he assumed. Within the confines discussed earlier, high- and low-anxiety individuals differ in their use of preattentive and attentional processes when presented with threatening stimuli, as well as in their use of interpretative mechanisms in the presence of ambiguous stimuli that can be interpreted in a threatening fashion.

The notion of trait anxiety as a semipermanent predisposition to experience anxiety that is largely determined by heredity implies a rather rigid conceptualization. The fact that personality changes over time, and is affected by environmental factors, means that a more dynamic model is required. It is, of course, a fundamental characteristic of the cognitive system that it has great capacity for change, and so at least some of the dynamic changes in trait anxiety over time can plausibly be attributed to alternations in cognitive processes and the contents of long-term memory.

Most physiologically based theories seem to imply a unidimensional view of trait anxiety. The expectation is that those individuals who are high in trait anxiety and thus have very responsive physiological systems should tend to be highly anxious in virtually all stressful situations. In fact, there is accumulating evidence for a multidimensional view of trait anxiety. Endler (1983) identified five different dimensions of trait anxiety, and reviewed studies in which the increase in state anxiety produced by a threatening or stressful situation as compared with a neutral situation was greater among those high in trait anxiety than among those low in trait anxiety only when there was congruence between the nature of the threatening situation and the dimension of trait anxiety being considered. The evidence provides especially strong support for a distinction between social evaluation and physical danger dimensions of trait anxiety.

The multidimensional nature of trait anxiety probably depends in large measure on the cognitive system. It may well be that individuals construct and develop different schemata or broad memory structures for each of the major areas of life (e.g., family; work; finance). The amount of anxiety experienced as a result of events in each area would then be influenced by the nature of the relevant schemata.

In sum, it is of interest that all of the major weaknesses with physiologically based theories of trait anxiety could potentially be substantially reduced by a theoretical formulation that included both the physiological and cognitive systems. In a sense, the cognitive system provides a gateway to the relevant physiological system (i.e., the visceral brain or BIS), and thus plays an important role in determining the nature of the physiological response. Of course, although this view emphasizes the effects of the cognitive system on the physiological system, there are undoubtedly many (relatively unexplored) effects of the physiological system on the cognitive system.

The difference in cognitive functioning between individuals high and low in trait anxiety affect what is processed and attended to and how it is interpreted. Those high in trait anxiety tend to have a preattentive selective bias that leads them to allocate processing resources to threatening stimuli, and they also tend to interpret ambiguous stimuli in a threatening fashion. Both of these tendencies increase the extent to which the environment is perceived as threatening, and thus may predispose to anxiety neurosis. In addition, the threat-related schemata or cognitive structures of those high in trait anxiety may also increase their vulnerability.

These various differences in cognitive functioning between individuals high and low in trait anxiety may very well be interrelated. For example, the preferred interpretation of an ambiguous stimulus may depend on which schema in long-term memory provides the best fit to the stimulus. The selective allocation of processing resources may also depend importantly on information stored in long-term memory. It remains an important task for the future to determine more precisely the ways in which cognitive structures and processes combine in individuals high and low in trait anxiety.

How might these various differences in cognitive functioning be incorporated into a clinically relevant cognitive model of trait anxiety? This issue (or at least closely related ones) is discussed more fully in the next chapter, but some preliminary considerations will be dealt with here. The kind of model that is appropriate may well resemble in some ways the cognitive model of depression put forward by Beck, Rush, Shaw, and Emery (1979). They argued that certain individuals possess maladaptive schemata or cognitive structures that predispose them to develop depressive symptoms. Clark and Hemsley (1985) have recently provided evidence that individuals high in trait anxiety tend to have frequent and intense depressive and anxious cognitions, and these cognitions may reflect underlying schemata that predispose to mood disorders.

According to Beck *et al.* (1979), schemata incorporating irrational beliefs or depressogenic assumptions may produce clinical depression when they are activated by relevant life events. The notion that clinical symptoms may result from the interactive influences of life events and schemata or cognitive structures is likely to be applicable to anxiety (cf., Beck & Emery, 1985), but even in connection with depression there is very little relevant research. An exception is an interesting recent study by Hammen, Marks, Mayol, and deMayo (1985). Their data broadly supported the view that negative events that are consistent with an individual's schemata will increase depression, whereas inconsistent negative events will not. Therefore, the impact of negative life events on an individual's emotional state can only be measured accurately provided that we take account of that person's cognitive structures or schemata. It is inadequate simply to consider the total number of negative life events, because that procedure assigns equal weight to schema-consistent and schema-inconsistent events.

We are in general agreement with the schema-based approach favored by Beck *et al.* (1979) and by Hammen *et al.* (1985), but their reliance on introspective evidence to assess schemata seems misplaced. For example, it appears that the irrational beliefs assessed introspectively may be transient concomitants of depression rather than the long-term vulnerability factor often assumed (Persons & Rao, 1985). In addition, there is the likelihood that much of the information contained in schemata is not accessible to consciousness. Our distinct preference is to assess schemata and other aspects of the cognitive system by means of laboratory tasks not dependent on introspection.

In sum, those high and low in trait anxiety differ in their preattentive selective biases, attentional processes, interpretative processes, and the contents of long-term memory at both molecular (e.g., specific worries) and molar (e.g., schemata) levels.

It is assumed that these aspects of cognitive functioning affect an individual's vulnerability to clinical anxiety, and so can be regarded as predisposing factors. Of course, the cognitive system interacts with other systems, and these interactions will ultimately have to be taken into account. Nevertheless, the cognitive system makes a distinctive contribution to individual differences in trait anxiety, as we have endeavored to show in this chapter.

TOWARD A COGNITIVE MODEL

This chapter has made it clear that there are numerous differences in cognitive functioning between those high and low in trait anxiety. In order to proceed to a cognitive model of trait anxiety that has potential clinical relevance, at least two goals have to be achieved: (a) it must be demonstrated that physiologically-based theories of trait anxiety are inadequate, and (b) aspects of cognitive functioning that predispose to anxiety neurosis must be identified.

Physiologically based theories of trait anxiety or neuroticism have been proposed by Eysenck (1967) and by Gray (1982), both of whom argued that there are genetically determined individual differences in the responsiveness of some physiological system. For Eysenck (1967), that system was the visceral brain, by which he meant "the hippocampus, amygdala, cingulum, septum, and hypothalamus" (p. 230). Gray (1982), as was mentioned earlier, argued for the importance of the behavioral inhibition system (BIS).

Even if the major contentions of these theories are supported experimentally, they cannot be regarded as complete accounts of trait anxiety. For example, it is presumably the case that the visceral brain or BIS will be activated primarily after stimuli have been categorized as threatening. In view of the accumulating evidence of individual differences in the initial processing of threatening stimuli, it is not usually possible to predict the extent to which a threatening stimulus will activate a given individual's physiological system in the absence of knowledge about the functioning of that individual's cognitive system.

ACKNOWLEDGMENTS

The ideas expressed in this chapter owe much to useful discussions with Colin MacLeod, but any deficiencies that remain are the responsibility of the authors.

REFERENCES

- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation*, Vol. 8 (pp. 47-89). London: Academic Press.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Beck, A. T., & Emery, G. (1985). *Anxiety disorders and phobias: A cognitive perspective*. New York: Basic Books.

- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Guilford Press.
- Blaylock, B. A. H. (1963). *Repression-sensitization, word association responses, and incidental recall*. Unpublished Master's thesis, University of Texas, Austin, TX.
- Borkovec, T. D., Robinson, E., Prujinsky, T., & DePree, J. A. (1983). Preliminary exploration of worry: Some characteristics and processes. *Behaviour Research and Therapy*, 21, 9–16.
- Bower, G. H. (1981). Mood and memory. *American Psychologist*, 36, 129–148.
- Bower, G. H., & Mayer, J. D. (1985). Failure to replicate mood-dependent retrieval. *Bulletin of the Psychonomic Society*, 23, 39–42.
- Bower, G. H., Gilligan, S. G., & Monteiro, K. P. (1981). Selectivity of learning caused by affective states. *Journal of Experimental Psychology: General*, 110, 451–473.
- Bransford, J. D. (1979). *Human cognition: Learning, understanding and remembering*. Belmont, CA: Wadsworth.
- Butler, G., & Mathews, A. (1983). Cognitive processes in anxiety. *Advances in Behaviour Research and Therapy*, 5, 51–62.
- Byrne, D. (1961). The Repression-Sensitization Scale: Rationale, reliability, and validity. *Journal of Personality*, 29, 334–349.
- Byrne, D. (1964). Repression-sensitization as a dimension of personality. In B. A. Maher (Ed.), *Progress in experimental personality research*. (Vol. 1, pp. 169–220). New York: Academic Press.
- Calvo, M. G. (1985). Effort, aversive representations and performance in text anxiety. *Personality and Individual Differences*, 6, 563–571.
- Carroll, D. (1972). Repression-sensitization and duration of visual attention. *Perceptual and Motor Skills*, 34, 949–950.
- Clark, D. A., & Hemsley, D. R. (1985). Individual differences in the experience of depressive and anxious, intrusive thoughts. *Behaviour Research and Therapy*, 23, 625–633.
- Coltheart, V., & Evans, J. St. B. T. (1981). An investigation of semantic memory in individuals. *Memory & Cognition*, 9, 524–532.
- Conley, J. J. (1984). The hierarchy of consistency: A review and model of longitudinal findings on adult individual differences in intelligence, personality and self-opinion. *Personality and Individual Differences*, 5, 11–25.
- Craske, M. G., & Craig, K. D. (1984). Musical performance anxiety: The three-systems model and self-efficacy theory. *Behaviour Research and Therapy*, 22, 267–280.
- Dixon, N. F. (1981). *Preconscious processing*. Chichester: Wiley.
- Dornic, S., & Fernaeus, S.-E. (1981). Individual differences in high-load tasks: The effect of verbal distraction. *Reports from the Department of Psychology, University of Stockholm*, No. 569.
- Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, 66, 183–201.
- Endler, N. S. (1983). Interactionism: A personality model but not yet a theory. In M. M. Page (Ed.), *Nebraska symposium on motivation*. New York: Plenum Press.
- Eysenck, H. J. (1967). *The biological basis of personality*. Springfield, IL: Charles C Thomas.
- Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences*. New York: Plenum Press.
- Eysenck, M. W. (1979). Anxiety, learning, and memory: A reconceptualization. *Journal of Research in Personality*, 13, 363–385.
- Eysenck, M. W. (1982). *Attention and arousal: Cognition and performance*. Berlin: Springer.
- Eysenck, M. W. (1984a). *A handbook of cognitive psychology*. London: Erlbaum.
- Eysenck, M. W. (1984b). Anxiety and the worry process. *Bulletin of the Psychonomic Society*, 22, 545–548.
- Eysenck, M. W. (1985). Anxiety and cognitive-task performance. *Personality and Individual Differences*, 6, 579–586.
- Eysenck, M. W. (in press-a). Individual differences in anxiety, cognition, and coping. In R. Hockey, A. Gaillard, & M. Coles (Eds.), *Energetics and human information processing*. Dordrecht: M. Nijhoff.
- Eysenck, M. W. (in press-b). Personality, stress arousal, and cognitive processes in stress transactions. In R. W. J. Neufeld (Ed.), *Advances in investigation of psychological stress*. New York: Wiley.
- Eysenck, M. W., MacLeod, C., & Mathews, A. (in press). Cognitive functioning and anxiety. *Psychological Research*.
- Floderus-Myrhed, B., Pedersen, N., & Rasmusson, I. (1980). Assessment of heritability for personality based on a short form of the Eysenck Personality Inventory. *Behavior Genetics*, 10, 153–162.

- Gray, J. A. (1982). *The neuropsychology of anxiety*. Oxford: Clarendon.
- Gray, J. A. (1985). Issues in the neuropsychology of anxiety. In A. H. Tuma & J. Maser (Eds.), *Anxiety and the anxiety disorders*. (pp. 5–25). London: Erlbaum.
- Gudjonsson, G. H. (1981). Self-reported emotional disturbance and its relation to electrodermal reactivity, defensiveness and trait anxiety. *Personality and Individual Differences*, 2, 47–52.
- Hamilton, V. (1983). *The cognitive structures and processes of human motivation and personality*. Chichester: Wiley.
- Hammen, C., Marks, T., Mayol, A., & deMayo, R. (1985). Depressive self-schemas, life stress, and vulnerability to depression. *Journal of Abnormal Psychology*, 94, 308–319.
- Haney, J. N. (1973). Approach-avoidance reactions by repressors and sensitizers to ambiguity in a structured free-association task. *Psychological Reports*, 33, 97–98.
- Ingham, J. G. (1966). Changes in MPI scores in neurotic patients: A three year follow-up. *British Journal of Psychiatry*, 112, 931–939.
- Klinger, E. (1985). A consciousness-sampling analysis of test anxiety and performance. *Journal of Personality and Social Psychology*, 47, 1376–1390.
- Krohne, H. W., & Rogner, J. (1982). Repression-sensitization as a central construct in coping research. In H. W. Krohne & L. Laux (Eds.), *Achievement, stress, and anxiety*. (pp. 167–193). London: McGraw-Hill.
- Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt, Brace, & World.
- Korhchin, S. (1964). Anxiety and cognition. In C. Scheerer (Ed.), *Cognition: Theory, research, promise*. (pp. 58–78). New York: Harper & Row.
- Lang, P. (1971). The application of psychophysiological methods to the study of psychotherapy and behavior modification. In A. Bergin & S. Garfield (Eds.), *Handbook of psychotherapy and behavior change*. (pp. 476–503). New York: Wiley.
- Leon, M. R., & Revelle, W. (1985). Effects of anxiety on analogical reasoning: A test of three theoretical models. *Journal of Personality and Social Psychology*, 49, 1302–1315.
- Lewinsohn, P. M., Berquist, W. H., & Brejle, T. (1972). The repression-sensitization dimension and emotional response to stimuli. *Psychological Reports*, 31, 707–716.
- Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, 20, 975–978.
- MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, 95, 15–20.
- Martin, M., Ward, J. C., & Clark, D. M. (1983). Neuroticism and the recall of positive and negative personality information. *Behaviour Research and Therapy*, 21, 495–503.
- Mayer, R. E. (1977). Problem-solving performance with task overload: Effects of self-pacing and trait anxiety. *Bulletin of the Psychonomic Society*, 9, 283–286.
- Mayo, P. R. (1983). Personality traits and the retrieval of positive and negative memories. *Personality and Individual Differences*, 4, 465–472.
- McKeon, J., Roa, B., & Mann, A. (1984). Life events and personality trait in obsessive-compulsive neurosis. *British Journal of Psychiatry*, 144, 185–189.
- Mogg, K., Mathews, A., & Weinman, J. (in press). Memory bias in clinical anxiety. *Journal of Abnormal Psychology*.
- Morris, L. W., Davis, M. A., & Hutching, C. H. (1981). Cognitive and emotional components of anxiety: Literature review and a revised worry-emotionality scale. *Journal of Educational Psychology*, 73, 541–555.
- Persons, J. B., & Rao, P. A. (1985). Longitudinal study of cognitions, life events, and depression in psychiatric inpatients. *Journal of Abnormal Psychology*, 94, 51–63.
- Schwartz, G. E., Davidson, R. J., & Goleman, D. J. (1978). Patterning of cognitive and somatic processes in the self-regulation of anxiety: Effects of meditation versus exercise. *Psychosomatic Medicine*, 40, 321–328.
- Simpson, G. B. (1984). Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin*, 96, 316–340.
- Spielberger, C. D., Gorsuch, R., & Lushene, R. (1970). *The state trait anxiety inventory (STAI) test manual*. Palo Alto, CA: Consulting Psychologists Press.
- Ullmann, L. P. (1962). An empirically derived MMPI scale which measures facilitation-inhibition of recognition of threatening stimuli. *Journal of Clinical Psychology*, 18, 127–132.

- Van Egeren, L. (1968). Repression and sensitization: Sensitivity and recognition criteria. *Journal of Experimental Research in Personality*, 3, 1-8.
- Wachtel, P. L. (1967). Conceptions of broad and narrow attention. *Psychological Bulletin*, 68, 417-429.
- Wagstaff, G. F. (1974). The effects of repression-sensitization on a brightness scaling measure of perceptual defense. *British Journal of Psychology*, 65, 395-401.
- Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin*, 96, 465-490.
- Weinberg, R. S., & Hunt, V. (1976). The interrelationships between anxiety, motor performance, and electromyography. *Journal of Motor Behavior*, 8, 219-224.
- Weinberger, D. A., Schwartz, G. E., & Davidson, R. J. (1979). Low-anxious, high-anxious, and repressive coping styles: Psychometric patterns and behavioral and physiological responses to stress. *Journal of Abnormal Psychology*, 88, 369-380.
- Williams, D. G. (1981). Personality and mood: State-trait relationships. *Personality and Individual Differences*, 2, 303-309.
- Young, G. C. D., & Martin, M. (1981). Processing of information about self by neurotics. *British Journal of Clinical Psychology*, 20, 205-212.

CHAPTER 10

Clinical Anxiety and Cognition

Andrew Mathews and Michael W. Eysenck

INTRODUCTION

In the preceding chapter we argued that clinical anxiety may be related to variations in what has been termed trait anxiety. Just as trait anxiety is described as the propensity to experience greater or lesser degrees of fear and anxiety, so particularly high levels of trait anxiety may predispose an individual to develop pathological anxiety states, given appropriate circumstances. In the present chapter we will therefore adopt the view that a satisfactory account of clinical states should be consistent with the theoretical model of trait anxiety discussed previously.

Such a position leaves open the question of whether clinical anxiety simply represents the upper end of a continuous dimension of normal anxiety, or whether some additional qualitative difference exists between anxiety disorders and normal emotional experience. The latter view may be suggested by clinical phenomena, such as apparently unprovoked panic attacks, which are seen rarely, if ever, in the normal population. Such apparently unique phenomena are not in fact incompatible with a fundamental underlying continuity between normal variations in anxiety and clinical conditions. In the psychophysiological model proposed by Lader and Mathews (1968) for example, higher anxiety is associated with increased levels of autonomic arousal, and decreased rate of habituation to repeated stimuli. As anxiety progressively increases, a point could be reached at which successive environmental stimuli arrive before the arousal caused by the preceding stimulus can dissipate. Under these circumstances autonomic arousal might spiral upwards, resulting in a panic attack.

Models of clinical anxiety such as those proposed by Eysenck (1967) and Gray (1982) also focus on the neurophysiological substrates of normal emotional behavior.

ANDREW MATHEWS • Department of Psychology, St. George's Hospital Medical School, Cranmer Terrace, London SW 17, England. MICHAEL W. EYSENCK • Department of Psychology, Birkbeck College, University of London, London WC1E 7HX, England. Some of the work reported here was supported by a grant from the Medical Research Council, UK.

In Gray's view, fear or anxiety is said to arise from the activity of the behavioral inhibition system, (BIS), located in the septo-hippocampal area of the mid-brain, but having excitatory input from centers in the brain stem. These lower centers set the sensitivity level of the BIS to specific classes of triggering stimuli; namely signals of punishment, frustrative nonreward, or novelty. Phobic states are handled within this model by the suggestion that certain classes of stimuli may have the innate capacity to trigger the BIS. More generalized anxiety states, where there is a less clearly defined external focus, are interpreted as arising from an elevated sensitivity of the BIS to all triggering stimuli. It is not clear from Gray's theory how such variations in sensitivity might arise. It is also unclear why different individuals develop anxiety states in response to different types of threatening events or circumstances. Such individual differences suggest the need to extend neuropsychological models of this sort to allow a role for life events, conditioning, and symbolic learning, all of which may modify the significance of specific threat stimuli for the individual. Just as we have argued that account must be taken of the cognitive system in understanding trait anxiety, so the same argument leads us to believe that cognitive processes play a role in the anxiety disorders.

THE ORIGINS OF PHOBIAS AND ANXIETY STATES

Classical conditioning of fear to biologically prepared stimuli (Seligman, 1971) may provide an explanation for the development of some specific phobias. However, phobias sometimes develop in the absence of direct aversive contact with the subsequently feared stimulus, and to stimuli that cannot by any stretch of the imagination be considered biologically prepared. Thus Rachman and Seligman (1976) have described cases of phobias of such unlikely stimuli as plants and chocolate. Rachman (1977) has gone on to argue that some fears may develop from vicarious learning experiences, or via information transmission from others. Recent evidence from animal experiments provides powerful support for the belief that modeling influences may be crucial in the development of some phobias. Rhesus monkeys typically acquire a severe and persistent fear of snakes following relatively brief observation of other monkeys showing fear reactions in the presence of a snake (Mineka, 1985; see also Chap. 4).

In the case of less specific fears, such as agoraphobia, it is more difficult to identify any relevant conditioning or modeling influences. However, a majority of agoraphobics do report that the disorder began with one or more acute attacks of panic, occurring at a time of increased background stress arising from a variety of sources (Buglass, Clarke, Henderson, Krietman, & Presley, 1977). Current psychological theories of agoraphobia thus suggest that anxiety attacks that are partly attributable to background stress are misattributed to external stimuli (Goldstein & Chambless, 1978; Mathews, Gelder, & Johnston, 1981). Avoidance may thus arise from the expectation that going out will provoke another panic attack, with dangerous or humiliating consequences. Secondary fears of collapse, physical illness or loss of control are commonly reported, and may serve to maintain avoidance behavior.

In so-called generalized anxiety and panic disorders, anxiety is often not attributed to any identifiable external stimulus or situation, although internal cues may

play an important role. An example is provided by one woman, diagnosed as suffering from generalized anxiety with hypochondriacal fears, in whom the relevant cue was a sore throat. The disorder developed about 2 years after she had nursed her father until his death from throat cancer. Some time later she experienced a persistent sore throat and attended her doctor repeatedly, until he eventually asked her if she was afraid that she too might have cancer. At this she became tearful and shortly afterwards developed a severe anxiety state. Despite reassurance from her doctor, awareness of any bodily sensation (particularly in her throat) triggered intense fear, with increasing conviction that she had cancer or some other serious disease.

In a systematic study of life events preceding anxiety or depressive disorders, Finlay-Jones and Brown (1981) found that depression tended to follow loss (such as a bereavement), whereas anxiety was more likely to follow so-called danger events (such as an illness that may or may not be serious). However, a substantial proportion of the sample studied did not develop anxiety disorders following even fairly severe threats, suggesting that other vulnerability factors were involved. At least two modifying influences suggest themselves; the prior vulnerability of the individual concerned, and the extent to which an event matches the specific concerns or worries of that person. As with depression (cf. Hammen, Marks, Mayol, & deMayo, 1985), an event will have greater impact if it is consistent with the existing concerns of an individual. Thus in the case history above, the upsetting experience of seeing a close relative die of cancer will presumably have created (or elaborated) a schema in memory concerned with the threat of serious disease. Such danger schemata may leave an individual particularly vulnerable to the impact of subsequent events that activate them.

There are also indications that the same frequency of events will have a greater impact on those with higher levels of trait anxiety. Obsessive-compulsive patients who were rated as having a highly anxious premorbid personality, experienced only half as many life events on average as did those with low premorbid anxiety, during the year prior to onset of the disorder (McKeon, Roa, & Mann, 1984). By implication, fewer events are necessary to precipitate a disorder in highly anxious individuals because they are more vulnerable to stress, whereas conversely a high density or severity of events is necessary to produce an anxiety disorder in those with low trait anxiety. Even previously very stable people can develop anxiety disorders when exposed to extremely traumatic threats, such as the continuous threat of death on the battle field or in concentration camps (Eitinger, 1964; Von Baeyer, 1969).

A COGNITIVE MODEL OF CLINICAL ANXIETY

Before reviewing the evidence supporting the view that clinical anxiety states are associated with characteristic cognitive processes (and possibly structures), we will outline the general model of emotion that has guided our thinking. We do not propose that this cognitive model should be seen as an alternative to the more physiologically oriented theories advanced by Eysenck (1967) or Gray (1982). Rather we envisage several complex systems operating at different levels in parallel, and that influence each other at various points.

It is assumed that emotional responses normally begin with the evaluation of incoming information having relevance to the goals or expectations of that individual. Because this applies to species other than man, we assume that semantic (language-based) processing is not necessary for this fundamental evaluative response (Martin & Levey, 1985), although such semantic processing may indeed influence some forms of emotional evaluation in humans. Presumably, the evaluative system is innately prepared to react with certain emotional responses to some biologically important stimuli. Thus heights may elicit fear and avoidance from infants prior to any opportunity for learning. Species that have adequate locomotion from birth avoid a "visual cliff" immediately, and show distress if placed above it (Marks, 1969, p. 21).

Emotional responding to stimuli can apparently also be acquired or modified without awareness that such a change has taken place. Repeated subliminal presentations of visual stimuli can lead to progressive increases in liking for those stimuli that had been exposed more frequently (the "mere exposure" effect) despite the fact that subjects could not recognize which had been shown previously (Kunst-Wilson & Zajonc 1980; Seamon, March, & Brody, 1984). Similarly, Bargh and Pietromonaco (1982) demonstrated that information presented outside awareness can influence subsequent social judgments about another person. Although it could be argued that such ratings are not in themselves evidence of an emotional state, findings of this type suggest it is unlikely that emotion arises only from conscious cognitive processing of events. Although Zajonc and Markus (1984) argue that these findings imply that emotion is noncognitive, such an argument seems to be based on an inappropriate identification of cognition with consciousness. Rather, we assume that at least some of the cognitive processes that are capable of generating emotional states are not available to consciousness.

Incoming information may thus trigger evaluative or emotional responses directly, as a result of innate preparation, novelty, or classical conditioning (Gray, 1982; Martin & Levey, 1985); or it may do so indirectly, after being processed for associative meaning at a higher level in the cognitive system. Thus the discovery of a small lump in one's body is likely to produce anxiety only if the possibility of cancer is thought about, and the extent of anxiety experienced will be related to knowledge of severity of the disease, effectiveness of treatment, and so on. It seems likely that after such higher-level processing of an event has led to an emotional reaction, subsequent emotional responding to the same or related events will become more rapid and automatic. That is, the evaluative system could contain representations of some events that are predetermined as emotional, and others that have been acquired and come to elicit emotion as a result of learning or symbolic processes. The way in which these acquired representations are organized, and the way in which they influence the processing of other information, forms the basis for our cognitive model of anxiety. Although it is not our intention to speculate about the neurophysiological mechanisms that could provide a basis for such processes, the evaluative system described here appears to overlap in function with that attributed by Gray (1982) to the BIS. Similarly the suggestion that higher-level processes are able to access and activate the emotional system (and vice-versa) means that these proposals are compatible in broad terms with cognitive models of emotion based on memory networks (e.g., Bower, 1981).

We can now apply this framework to clinical anxiety by proposing that vulnerability to anxiety disorders may arise as a result of systematic biases in the cognitive input to the emotional evaluative system. As discussed in the previous chapter dealing with trait anxiety, such biases may occur at a number of processing stages. Given that some of the information in most environments may be seen as potentially threatening, a preattentive bias could result in the more threatening aspects being successful in capturing processing resources. Once captured, selective attention would ensure the additional intake of information from sources of threat, whereas sources of more reassuring information are ignored. This in turn might lead to the preferential encoding into memory of information, which has the capacity to rearouse anxiety when activated subsequently. Finally, more elaborated or more accessible schemata in memory concerning danger might lead to a bias in judgment of ambiguous situations, such as the overestimation of personal risk (Butler & Mathews, 1983).

These biases cannot, of course, be considered independently of biological differences, such as autonomic arousability, or of life event frequency. Variations in the experience of fear and anxiety as a result of innate and environmental influences will inevitably be stored in memory, and thus modify the cognitive representations that we have termed danger schemata. However, we would maintain that the causes of an anxiety disorder can only be fully characterized by taking into account how these cognitive representations are organized and influence the future processing of threatening information.

COGNITIVE CONTENT IN ANXIETY DISORDER

The views previously expressed clearly owe much to Beck and his colleagues (cf. Beck & Emery, 1985; Beck & Rush, 1975). As indicated earlier, however, we differ in the reliance that should be placed on introspective evidence and reported thought content. Beck, Laude, and Bohnert (1974) for example, used as basic data the reports on patients with generalized anxiety disorders about their thoughts at times of increased anxiety. In all cases thought content was said to involve the theme of personal danger, although in some cases this was predominantly physical (e.g., death, disease, accident, or assault) and in other cases it was more social in nature (e.g., failure, rejection, humiliation).

In a replication of these observations, Hibbert (1984) adopted a somewhat more standardized form of questioning, and again concluded that reported ideation involved physical or social dangers. After dividing his sample into those with and without panic attacks, Hibbert suggested that panics tended to be associated with more catastrophic thought content, particularly that concerned with sudden death or severe disease. For example, one patient with panics reported the thought that he might die suddenly of a heart attack. This thought was accompanied by subjective difficulty in breathing and heart palpitations, which began following his father's death during cardiac surgery. Another patient, who did not have panic attacks, reported the thought that she might make a fool of herself in public. This thought was accompanied by a muscular tremor and there was no clear precipitating event other than marital conflict. If substantiated by further research, these findings may

indicate that panics arise from a vicious circle of influence between thoughts of imminent collapse or serious disease, and physiological fear reactions that generate yet more fear when they are noticed.

In contrasting cognitive content associated with generalized anxiety and phobic states, Beck and Rush (1975) conclude that whereas both concern personal danger, thoughts reported by phobics were restricted to relatively specific external sources of danger, and were less often concerned with internal body sensations or the thought processes themselves. Thus it is possible to argue that the content of thoughts may influence the type of anxiety disorder experienced, with specific externally directed focus leading to phobias, thoughts of psychological or social failure being associated with generalized anxiety, and panic disorders being attributable to catastrophic thoughts of physical disaster. In practice, the relationship between thought content and type of disorder is frequently not as precise as this argument implies.

There are many difficulties about the interpretation of such self-reported introspective evidence. Nonanxious control groups have not been studied, and even if such controls were to be included, demand or expectancy effects would offer obvious rival explanations for any differences. Despite these criticisms, we regard descriptive work of this kind as useful in generating hypotheses that can be tested more rigorously in later experiments. It may be possible, for example, to construct hypotheses about the way in which the contents of long-term memory are organized by analyzing the structure of self-reported worries in anxious and nonanxious samples.

Using the worry questionnaire cited in the previous chapter, we have examined results from a small sample of anxiety state patients, and found that relative to a control group they report increased frequency of worry in almost all the clusters identified (coping with responsibility, intimate relationships, social rejection, personal fulfilment, financial troubles, and physical danger). The sample was too small to run a separate cluster analysis of the anxiety group alone, so that it is not known whether they might show an idiosyncratic structure. However, analysis of variance using scores from the clusters identified revealed a significant interaction between type of worry and subject group, due to equivalent scores being found for only one cluster: general societal and environmental problems. That is, generalized anxiety states were associated with increases across the whole range of worries, with the exception of those that did not refer directly to oneself. This finding, although very tentative, may indicate that the results are not entirely attributable to demand effects, and is consistent with other evidence suggesting a self-focus in cognitions concerned with danger in highly anxious individuals (e.g., Butler & Mathews, 1983). As indicated earlier however, we do not regard such self-report measures as necessarily valid indicators of underlying process or structure. For this reason we turn now to the experimental study of cognitive processes using objective measures.

ATTENTION AND CLINICAL ANXIETY

There is some evidence that phobic and obsessional patients show greater vigilance for cues related to their specific concerns than they do for unrelated information. Burgess, Jones, Robertson, Radcliffe, and Emerson (1981) reported that

phobics were better able than were controls to detect words related to their phobias in the nonattended channel during a dichotic listening experiment. Such an effect may be attributable to the greater frequency of usage made of phobic-related words by these subjects, because more frequently used words are generally easier to detect. However, in reporting a very similar phenomenon in obsessive-compulsive subjects, Foa and McNally (1986) argued that the reduction of interference found following treatment showed that the emotional significance of the words was of greater importance. Because treatment involved exposure to related material, increased familiarity alone seems unlikely to be responsible.

The argument that the emotional significance of the words used determines the extent to which vigilance is shown, is strengthened by parallel findings in nonclinical populations. Mothers of children about to have an operation showed increased detection rates for words embedded in a background of noise when the words related to surgery (Parkinson & Rachman, 1981). Despite this evidence that current concerns influence detection rate, it is not clear that selective attention is necessarily involved. Because subjects in these experiments were required to detect and report the words they heard, a response bias explanation is equally plausible. On the basis of a similar amount of partial information, subjects may be more likely to guess within an area that is of current concern to them.

Studies that have examined the disruptive effects of threat cues that are incidentally presented during unrelated tasks are not so vulnerable to a response bias explanation. Watts, McKenna, Sharrock, and Trezise (1986) have shown that relative to controls, spider phobics are slower to color-name words related to spiders (e.g., web) despite instructions to ignore word content. As in the study by Foa and McNally (1986), Watts *et al.* found that the difference between phobic and control subjects disappeared following treatment, suggesting that emotional meaning rather than priming by recent exposure (Warren, 1972) was responsible.

Although the mechanisms underlying interference with color-naming are still ill-understood, this evidence suggests that emotionally significant words may cause difficulty due to subjects' inability to avoid attending to them. Presumably, all subjects automatically access the meaning of words they are required to color-name, but further processing is normally inhibited by the demands of the task. When the accessed meaning is of particular significance to that individual this may lead to further processing, or emotional reactions, which interfere with task performance.

In a related experiment, Mathews and MacLeod (1985) required subjects with generalized anxiety states to color-name sets of words that denoted either physical (e.g., cancer, mutilated) or social (e.g., inferior, failure) threat. Nonanxious controls were not slowed in color-naming performance by the presence of threat words, in comparison with positive words that were incompatible with danger. The anxious subjects on the other hand were significantly slower in the presence of both types of threat material than they were in color-naming the nonthreat words.

To determine whether the interference was related to reported cognitive content, anxious subjects were divided into those reporting predominantly physical concerns (e.g., disease or accident, etc.) and those whose concerns were more related to social threat (e.g., being criticized, or failing in some way). Only those anxious subjects with physical worries showed interference with physical threat words, although all

were affected by social threat words. Thus there was some evidence that interference was greatest when the interfering material matched the self-reported content of worries, giving partial support to the idea that danger schemata may be structured differently in different anxious subpopulations. Furthermore, because the magnitude of the effect correlated significantly with state anxiety score, mood state may be related to the level of activation in these cognitive schemata.

These data suggest that an interaction between current mood and domain of conscious concern results in selective processing of threat cues, although alternative explanations remain to be examined. In particular, it is unclear whether an automatic process is involved, or whether some difference in the strategies adopted by subject groups may be responsible for the observed interference. This question was addressed using a dichotic listening paradigm, in which subjects shadowed neutral stories while ignoring threat or nonthreat words in the unattended channel (Mathews & MacLeod, 1986). To detect the extent to which the unattended material captured processing resources, subjects were required to perform a simple visual detection task, synchronized with the threat or nonthreat words, so that increased processing demand would be revealed by longer detection latency. As expected, generally anxious subjects were slowed in their visual reaction times when probes coincided with threat as opposed to nonthreat words, whereas the performance of control subjects was not affected by the type of unattended material. This finding strengthens the hypothesis that anxious subjects tend to divert processing resources away from the on-going task in the presence of information related to threat. Because none of the subjects were able to report any of the unattended words afterwards, or to recognize them above chance level, this selective processing bias appears to be automatic, in the sense of being independent of awareness.

The alternative hypothesis, that additional processing resources were required to divert attention away from threat stimuli, seems to be definitively ruled out by the results of an experiment on the direction of attentional deployment (MacLeod, Mathews, & Tata, 1986). Generally anxious patients were required to read the upper word in a series of pairs, while ignoring the lower word, and to respond as rapidly as possible to a small dot that sometimes appeared in the space just vacated by either one of the words. The latency to detect probe dots that replaced different types of words was used to assess the location of attentional deployment. Regardless of upper (attended) or lower (unattended) location, anxious subjects responded more rapidly to probes that replaced threat rather than nonthreat words. It is thus clear that anxiety is associated with a shift in attentional resources toward information related to threat, rather than away from it. Interestingly, the opposite pattern of shifting away from threat words was found with the nonanxious controls. By implication, avoidance of mildly emotional threat cues at a very early stage of its processing is characteristic of nonanxious subjects, perhaps because it protects against repeated and unnecessary arousal. By contrast, individuals with anxiety disorders (and perhaps with high trait anxiety) seem unable to inhibit attentional shifts toward stimuli associated with threat, even when these stimuli are irrelevant to the task at hand. Although no physiological reactions were recorded in these studies, it is tempting to speculate that such attention shifts are accompanied by arousal increases, and that these cognitive phenomena may have some relationship to the psychophysiological characteristics of anxiety states described by Lader and Mathews (1968).

However, two negative findings complicate interpretation of the findings on attentional deployment. One is the failure in two out of three related experiments (Mathews & MacLeod, 1986; MacLeod *et al.*, 1986) to find significant correlations between the size of the effect studied and either state or trait anxiety level, contrary to earlier suggestions (Mathews & MacLeod, 1985). Because there were clear differences between anxious and nonanxious groups in all three experiments, this suggests some kind of threshold effect in the relationship between attentional deployment and anxiety or the operation of some other interacting process that obscures any simple relationship.

The second complication is that the original suggestion of a match between type of material attracting attention and reported domain of concern (physical vs. social), also failed to replicate in the later experiments. Here the most plausible explanation may be that the early stages of attentional deployment involve only the simple categorization of material as potentially threatening or not, and further processing according to type of threat then occurs later. It may be this further processing that was responsible for slowed color-naming performance, because interference in the Stroop color-naming test can occur at early and late stages of processing (Stirling, 1979).

RECALL AND RECOGNITION IN CLINICAL ANXIETY

Given the findings on attentional deployment and the suggestions about differences in long-term memory structure between high- and low-trait-anxious subjects, it is to be expected that clinically anxious individuals will have a bias in memory favoring material related to threat. Such a bias should arise because information related to threat is attended to selectively, and would be easy to retrieve because of its encoding within danger schemata. Despite this expectation, evidence of bias in memory has been more difficult to obtain in clinical anxiety than in depression, for reasons that are not yet entirely clear.

Evidence that high trait anxiety (or neuroticism) is associated with a bias favoring recall of negative information about oneself, has been reviewed in the previous chapter (e.g., Mayo, 1983; Martin, Ward, & Clark, 1983). Similar evidence of bias has been reported with agoraphobic subjects, who were better able to recall words or passages related to their fears than were nonanxious controls tested with the same material (Nunn, Stevenson, & Whalan, 1984). This latter study is not easy to interpret, however, because the words and passages involved would have implied very different things to the two groups. An account of a trip to the supermarket, for example, obviously has very different connotations to an agoraphobic than to a nonanxious control subject.

In studies of memory bias in depressed populations, it is common to use both self and other encoding tasks, and to present adjectives having either a negative or positive connotation, prior to subsequent recall (e.g., Bradley & Mathews, 1983). Such studies have revealed that depressed subjects recall relatively fewer positive adjectives, but only when these have been encoded in relation to oneself. As a result, it is difficult to interpret the results as being due to individual word usage effects, or to a simple response bias favoring all negative material. Both of these interpretations may apply to results obtained by Nunn *et al.* (1984) or Mayo (1983), although

not to those of Martin *et al.* (1983), who found neuroticism in students was related to negative recall bias only for self-encoded words.

When essentially identical procedures have been used with generally anxious patients, however, results have been strikingly different. Mogg, Mathews, and Weinman (1987) required clinically anxious subjects and normal controls to judge whether positive, negative, or supposedly threatening words (e.g., humiliated) applied to themselves or to a well-known television personality. Not only was there no recall effect favoring negative or threatening words, but a signal detection analysis of subsequent recognition revealed a trend favoring poorer discrimination of negative words by anxious subjects. Furthermore, in one of the studies discussed previously (Mathews & MacLeod, 1985), recognition tests failed to show any differences between groups even for the threat words that had previously produced differential interference effects. Finally, in an unpublished experiment using a different paradigm involving intentional recall of word triads (which included both threat and nonthreat words), we have found slight evidence of relatively poorer immediate recall for threat material in anxious subjects.

Although this evidence seems consistently negative, the apparently poorer recall of threat words by anxious subjects in the last experiment cited was confined to immediate recall, and tended to reverse after a delay. In the light of this finding, it may be prudent to await further data on long-term recall before arriving at any conclusion. The tendency for emotional material to be recalled poorly in the short term and better in the long term has been documented on several previous occasions (e.g., Parkin, Lewinsohn, & Folkard, 1982). Nonetheless, the pattern of results overall is in marked contrast with those found in the case of depressed populations.

Two possible explanations may be worth considering; the role of competition for attentional resources and of avoidance strategies that may be adopted by anxious subjects. In all the studies that have shown attentional bias effects there has been competition for processing resources between threat and nonthreat stimuli. When such competition does not exist (e.g., the lexical decision paradigm), preliminary data from our laboratory suggests that there are no differences between groups in their ease of identifying positive versus negative words presented singly. Typically, memory experiments do not present multiple stimuli that compete for attention, and so may be insensitive to differences based on attentional deployment. Second, recall differences may depend on secondary reactions and strategies adopted by anxious subjects, after they have identified material as threatening in content. Unlike depressives, anxious subjects may tend to adopt secondary cognitive avoidance strategies, such as selective ignoring or failure to rehearse, which could inhibit or delay elaborative encoding.

JUDGMENT AND INTERPRETATION IN CLINICAL ANXIETY

The evidence that trait anxiety is associated with a tendency to interpret ambiguous situations as threatening has been reviewed in the previous chapter. In the study reported by Eysenck, MacLeod, and Mathews (in press), both generalized

anxiety patients and high-trait control subjects tended to select the spelling of homophones (e.g., dye, die) that corresponded to the more threatening interpretation, compared with low-trait controls. In the clinical sample, the number of threatening interpretations selected correlated significantly with both trait (0.73, $p < .01$) and state (0.66, $p < .01$) anxiety level. Once again, however, there were no differences between anxious subjects who reported that they worried more about physical as opposed to social concerns, in terms of extent of bias shown with each type of material. If the hypothesis advanced previously is correct, this failure to find more specific effects may arise because attention is first drawn towards any threatening alternative, prior to more extensive processing concerning the precise nature of the threat.

When making judgments of more naturalistic events, generally anxious patients behave in a way that might be expected from the foregoing discussion. In estimating the probability of uncertain future events, anxious subjects rate the risk of negative items to be higher than do nonanxious controls, although the two groups do not differ in the case of positive events (Butler & Mathews, 1983). There was also a strong tendency for the anxious subjects to give higher probability ratings for future negative events that could happen to themselves, as opposed to the same events rated for someone else. By implication, their judgments were arrived at by accessing information relating to personal danger, rather than abstract knowledge of dangers in general. Subjective probability is often assumed to be influenced by the use of the availability heuristic; that is, the readiness with which examples of the class of event to be rated can be accessed (Kahneman, Slovic, & Tversky, 1982). It may be, therefore, that anxious individuals are faster to retrieve information about personally threatening events, because such events are stored in currently active schemata, and that this retrieval bias causes inflation of subjective risk. This idea has been partially supported by further work in progress, showing significant associations between areas of reported worry and the type of events that show relatively high subjective probability. That is, an individual who worries most about, say, being rejected by others, will also tend to rate the future risk of such rejection to be relatively high. On the other hand there have been some contrary findings, such as a failure to show any significant association between latency to recall or imagine examples of threatening events, and the subjective risk of such events.

These apparently contradictory findings may not be so irreconcilable as they first appear. Worry within a certain domain of concern may indeed have inflated subjective risk in the past, but only this previous estimate of risk needs to be recalled when subjects are asked to make a rating. Another related possibility is that inflation of subjective risk in anxiety is not based on the conscious retrieval of any specific events or estimates, but rather reflects the operation of a general rule arising from the repeated experience of attending to the more threatening aspects or interpretations of environmental events. Mild or ambiguous threat cues are everyday occurrences, provided one is actively alert to them. Minor bodily sensations may be the first sign of a serious illness, a friend's apparent preoccupation may signify disapproval, a stranger on a dark street could be a potential assailant, and so forth. Repeated awareness of such threats could result in a general mental set, or underlying assumption, that one is at risk from a range of dangers, without one necessarily being able to recall the specific instances that led up to such an assumption. It is even possible

that in generalized anxiety states, some individuals may never become aware of specific environmental cues that trigger anxiety, but only experience the feeling of being in danger. Either way it appears that the heightened subjective risk for negative events that results is more likely to occur at the time of encoding, rather than being dependent on subsequent retrieval.

CLINICAL IMPLICATIONS OF COGNITIVE RESEARCH

None of the research discussed so far conclusively demonstrates that biased cognitive processing either causes or contributes to the maintenance of pathological anxiety. It remains possible that clinical anxiety states arise in quite different ways, and the cognitive effects documented here are secondary consequences of the emotional disorder rather than being one of its causes. Unfortunately, it is difficult to examine causal direction without carrying out a large-scale prospective study of the general population, and testing the power of preexisting cognitive bias to predict breakdown under stress. Unless such a study can be done, the hypothesis that trait anxiety and the cognitive processes underlying it are vulnerability factors for clinical anxiety states will remain untested. At the very least however, data showing cognitive differences between anxious, depressed, and normal populations must be accounted for in any satisfactory theory of mood disorders.

In the case of depression there is some tentative evidence that cognitive distortions interact with negative mood in such a way as to influence the duration of the depressive episode (Lewinsohn, Steinmetz, Larson, & Franklin, 1981). Teasdale (1983) has drawn on this and other data to argue that mood state and negative recall bias influence one another in a circular fashion. Thus, a depressed individual will tend to access negative memories more readily, which will in turn serve to maintain depressed mood. Those who show the greatest recall bias in response to depressed mood will thus tend to remain depressed longer.

Such a circular model is also possible in anxiety states. Perhaps anxiety leads to increased attention being paid to threatening events, which in turn serves to maintain anxiety. Attentional rather than memory processes are implicated in anxiety, because the former have been amply demonstrated, whereas recall effects have not. Conversely, memory bias is easy to demonstrate in depression, but attentional effects are not (e.g., MacLeod *et al.*, 1986). It is therefore tempting to speculate that the two emotional disorders are associated with characteristic biases in different cognitive processes. Such a distinction would fit with the clinical observation that depressed patients tend to ruminate on past losses or personal failings, whereas anxious patients are more vigilant concerning future dangers that they are concerned to avoid.

Another method of testing causal direction is that of psychological intervention or treatment. If it could be shown that treatments based on cognitive theories lead to better outcome than did alternative methods, perhaps the theories could be seen as being strengthened. Unfortunately, there are a number of problems with this argument, not the least of which is the lack of data on cognitive treatments for clinical anxiety. Beck and his colleagues (e.g., Beck & Emery, 1985) propose that anxiety disorders should be treated by cognitive therapy modeled on the method they have

developed for use in depression. However, whereas there is fairly good evidence now that cognitive therapy is at least as effective for mild to moderate depression as antidepressant medication (e.g., Murphy, Simons, Wetzel, & Lustman, 1984), in the case of generalized anxiety no outcome data have yet been reported.

There are slight indications that therapy that combines cognitive components with exposure may improve long-term outcome in phobias (Butler, Cullington, Munby, Amies, & Gelder, 1984), although the more typical outcome is that exposure alone is as effective as the combination, and that exposure is clearly superior to cognitive therapy alone (e.g., Emmelkamp & Mersch, 1982). Perhaps the only convincing evidence that cognitive therapy can be as effective as behavioral techniques arises from studies of volunteers with performance anxieties (e.g., Kendrick, Craig, Lawson, & Davidson, 1982; Meichenbaum, Gilmore, & Fedoravicious, 1971). However, even if future research shows that verbal cognitive therapy is less effective than behavioral techniques, this will not necessarily reflect on the cognitive model developed here. The reason for this is that cognitive therapy as presently practiced may or may not be effective in changing the processes and biases postulated to cause anxiety. These underlying processes may not be consciously accessible, and might in any case be relatively impervious to verbal modification. This last point has been advocated by cognitive theorists, such as Bandura (1977), who none the less argue that the observation of one's own behavior and its consequences is the most powerful method of producing cognitive change.

A further problem with using the outcome of cognitive treatments as evidence for the validity of the theory is that little agreement exists on the definition and limits of cognitive as opposed to behavioral treatments. Traditionally, systematic desensitization in imagination has been considered to be a behavioral treatment, despite consisting largely of verbal instructions to imagine phobic scenes. It now seems extremely unlikely that desensitization depends on counterconditioning or extinction, so that little if any justification for categorizing it as a behavioral treatment remains. Relaxation is not necessary in order to achieve therapeutic effects, and there is no convincing evidence that change in physiological reactions to imaginal scenes generalizes to real stimuli (Mathews, 1971).

Most conclusively, therapeutic effects can be achieved using a convincing imitation of desensitization which purports to use subliminal stimuli (actually non-existent) and demonstrable progress in reducing physiological fear reactions (actually faked). In such an experiment, Lick (1975) found that identical behavioral improvement followed the dummy procedure and systematic desensitization in imagination. How can such results be explained? The most likely explanation seems to be that desensitization and the dummy procedure are quite effective methods of inducing cognitive change. In both cases, subjects are presented with apparent evidence from their own reactions that fear has declined. Although this does not necessarily change physiological responses to the real stimulus, it may well have a powerful motivational effect on subsequent approach behavior, and perhaps it is this behavioral change that leads to reductions in the other parameters of fear (Mathews, 1971; Mathews *et al.*, 1981).

It is generally accepted that treatments involving systematic exposure to feared situations in reality are more effective than when the same situations are presented

in imagination (e.g., Foa & Kozak, 1985). The same reviewers also concede that imaginal desensitization is more effective than less convincing alternatives, or (obviously) than no treatment at all. To the extent that desensitization is now regarded as a cognitive treatment, this provides some evidence for the role of cognitive factors in maintaining or reducing fear behavior. As to why desensitization is particularly effective in changing behavior, recent research suggests that, under some circumstances, imagery will lead to cognitive changes that resemble those acquired through actual experience.

When subjects are required to perform actions on some occasions, and only to imagine doing them on others, their memories of actually doing something are readily confused with memories of actions that have only been imagined (e.g., Anderson, 1984). In other studies (e.g., Gregory, Cialdini, & Carpenter, 1982) vividly imagining a particular future scenario was found to influence the judged probability of the event occurring in actuality. Thus, in certain respects, the cognitive system treats the content of induced imagery as having some similarities to evidence gathered from the real world. It is therefore possible, although by no means certain, that the capacity of imagery to modify the contents of long-term memory explains the effectiveness of desensitization.

In a series of psychophysiological studies, Lang (e.g., 1979, 1985) has developed a model of phobic imagery that relates to this issue. He suggests that imagery instructions, particularly when they include response as well as stimulus elements, are able to access the contents of phobic schemata in memory, and thus allow changes in cognitive structure. Evidence for this view includes the finding that individuals who report more vivid imagery, and those whose images are accompanied by greater physiological reactions, tend to improve more following systematic desensitization. Lang and his colleagues have also shown that phobic imagery induced with response-oriented instructions (e.g., "My heart is beating fast, I want to run away") elicits greater physiological reactions than do stimulus-oriented instructions alone. It remains to be shown whether or not such response-oriented instructions improve the outcome of desensitization, as Lang's theory would predict. Indirect support comes from a study by Borkovec and Sides (1979), who tested predictions made by Mathews (1971) that training in relaxation should enhance vividness and initial response to imagery, while maximizing subsequent decline. As expected, speech phobic subjects who relaxed synchronously with imagery induction showed higher heart rate reactions at first but these subsided more rapidly over trials. This in turn led to greater reductions in subjective fear (but not heart rate) during actual public speaking.

It is still not certain why this pattern of greater initial physiological response and subsequent decline leads to better outcome. To argue as does Lang that physiological reactions provide evidence that the relevant schemata has been accessed does not explain why its contents are changed by repeated evocation. At present the hypothesis most plausible to us is that this pattern of response is an index of the subject's engagement in the task, and thus the probability that information from the imagery instructions, and the subject's own response to it, will be encoded. Hence, Borkovec (1982) has suggested, on the basis of other work, that imaginal desensitization or flooding that incorporates instructions to imagine avoidance or failure to

cope results in less fear reduction, due to decreased functional exposure. We would extend this suggestion by adding that full involvement in the procedure will lead to evidence of reduced fear reactions to (imagined) phobic stimuli being accepted as valid. As a result, phobic subjects will suffer less from anticipatory anxiety, and be more motivated to enter the actual phobic situation. Because there is no evidence that relaxation during desensitization, or expectancy manipulations, alters physiological reactions during posttreatment behavioral testing (e.g., Borkovec & Sides, 1979), we can assume that the cognitive change during imaginal desensitization does not eliminate automatic emotional evaluations directly. Instead, the change is probably confined to consciously accessible propositions that relate to anticipatory anxiety or predictions about the extent to which one can control or tolerate fear in the real situation. Subsequent reduction of fear following repeated exposure to the real phobic stimulus may then come about through a number of mechanisms, such as habituation, extinction, or cognitive reappraisal.

It would be a disappointing outcome to the cognitive research developments that have been described earlier if they resulted only in the reinterpretation of existing treatments, without contributing new ideas for the assessment and treatment of anxiety. Although this may indeed be the case, there are perhaps some indications of the directions in which progress can be made concerning the treatment of generalized anxiety disorders. It is now clear that generalized anxiety is associated with selective processing of threat cues, even if this occurs in the absence of awareness. Such a conclusion implies that external (as well as internal) stimuli may be more relevant in this condition than was previously believed, perhaps because of the frequent inability of such patients to report on the relevant events. This should encourage therapists to search for triggering stimuli by direct observation rather than to rely on their clients' introspections after the event. It also raises the question of a greater role for exposure in the treatment of generalized anxiety. Many methods of anxiety management include exposure to anxiety-eliciting stimuli as a device to teach and practice control skills. Because there is some evidence suggesting that these methods are effective in clinical populations (e.g., Butler, Cullington, Hibbert, Klimes, & Gelder, *in press*; Jannoun, Oppenheimer, & Gelder, 1982), perhaps this exposure element is responsible for at least part of their effect.

Cognitive therapy (Beck & Emery, 1985) also includes an element of exposure, in that clients are often asked to test out their (presumably biased) expectations in real situations, in order to disconfirm them. Following the argument made earlier in discussing desensitization, to the effect that cognitive manipulations employing verbal or imaginal methods are relatively ineffective in modifying automatic emotional reactions, the real-life practice element of cognitive therapy may be more effective in this sense than are the verbal components. By the same token, the role of discussing thoughts and beliefs in cognitive therapy may be to facilitate exposure to anxiety-provoking stimuli, thus allowing change in the less consciously accessible cognitive processes underlying emotion. Such a line of reasoning also suggests that cognitive therapy could be strengthened by combining it more systematically with exposure to anxiety-evoking cues. Because cognitive therapy typically involves discussion of thoughts and beliefs under relatively calm conditions, state-dependent learning effects

(cf. Bower, 1981) will minimize transfer to real-life situations associated with anxiety. Exposure to anxiety-evoking cues during treatment sessions would reduce this problem, and at the same time facilitate change in cognitive processes not available for self-report.

In general terms, the future development of cognitive therapy will depend on a combination of clinical intuition, clinical practice, and the contribution of laboratory-based research and theory into cognitive functioning among anxious patients. At the very least, the delineation of salient differences in cognitive functioning between anxious patients and normal controls can provide a very valuable first step in the evolution of cognitive forms of therapy. Where laboratory findings confirm clinical intuition, they serve the useful purpose of strengthening the support for current cognitive approaches to therapy. Where laboratory findings (e.g., the discovery of preattentive selective biases) extend our knowledge of cognitive functioning in anxious states, they may serve to suggest new orientations for therapy.

To summarize, we have attempted to show that anxiety states are associated with characteristic cognitive processes that influence emotional and evaluative reactions. We have proposed that these cognitive processes can be seen as arising from the activity of knowledge structures in long-term memory, termed danger schemata, and result in the selective processing of information related to personal threat. This cognitive model is seen as complementing more biologically oriented views, and may thereby enrich our understanding of the etiology and treatment of anxiety states.

ACKNOWLEDGMENTS

Thanks are due to Tom Borkovec for helpful comments on an earlier draft of this chapter.

REFERENCES

- Anderson, R. E. (1984). Did I do it or did I only imagine doing it? *Journal of Experimental Psychology: General*, *113*, 594–613.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215.
- Bargh, J. A., & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology*, *43*, 437–449.
- Beck, A. T., & Emery, G. (1985). *Anxiety Disorders and Phobias: A cognitive perspective*. New York: Basic Books.
- Beck, A. T., & Rush, A. J. (1975). A cognitive model of anxiety formation and anxiety resolution. In I. Sarason & C. Spielberger (Eds.), *Stress and anxiety*. Vol. 2. (pp.69–80). New York: Halsted Press.
- Beck, A. T., Laude, R., & Bohnert, M. (1974). Ideational components of anxiety neurosis. *Archives of General Psychiatry*, *31*, 319–325.
- Borkovec, T. D. (1982). Facilitation and inhibition of functional CS exposure in the treatment of phobias. In J. Boulougouris (Ed.), *Learning theory approaches to psychiatry* (pp. 95–102). New York: Plenum Press.
- Borkovec, T. D., & Sides, J. K. (1979). The contribution of relaxation and expectancy to fear reduction via graded, imaginal exposure to feared stimuli. *Behaviour Research and Therapy*, *17*, 529–540.
- Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*, 129–148.

- Bradley, B., & Mathews, A. (1983). Negative self-schemata in clinical depression. *British Journal of Clinical Psychology*, 22, 173–181.
- Buglass, D., Clarke, J., Henderson, A. S., Krietman, N., & Presley, A. S. (1977). A study of agoraphobic housewives. *Psychological Medicine*, 7, 73–86.
- Burgess, I. S., Jones, L. M., Robertson, S. A., Radcliffe, W. N., & Emerson, E. (1981). The degree of control exerted by phobic and non-phobic verbal stimuli over the recognition behavior of phobic and non-phobic subjects. *Behaviour Research and Therapy*, 19, 233–243.
- Butler, G., & Mathews, A. (1983). Cognitive processes in anxiety. *Advances in Behaviour Research and Therapy*, 5, 51–62.
- Butler, G., Cullington, A., Munby, M., Amies, P., & Gelder, M. (1984). Exposure and anxiety management in the treatment of social phobia. *Journal of Consulting and Clinical Psychology*, 52, 642–650.
- Butler, G., Cullington, A., Hibbert, G., Klimes, I., & Gelder, M. (in press). Anxiety management for persistent generalized anxiety. *British Journal of Psychiatry*.
- Eitinger, L. (1964). *Concentration camp survivors*. Oslo: London.
- Emmelkamp, P. M. G., & Mersch, P. (1982). Cognitive and exposure in vivo in the treatment of agoraphobia. Short term and delayed effects. *Cognitive Therapy and Research*, 6, 77–88.
- Eysenck, H. J. (1967). *The biological basis of personality*. Springfield, IL: Charles C Thomas.
- Eysenck, M. W., MacLeod, C., & Mathews, A. (in press). Cognitive functioning and anxiety. *Psychological Research*.
- Finlay-Jones, R. A., & Brown, G. W. (1981). Types of stressful life and life events and the onset of anxiety depressive disorders. *Psychological Medicine*, 11, 803–815.
- Foa, E. B., & Kozak, M. J. (1985). Treatment of anxiety disorders: implications for psychopathology. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders*. (pp. 421–452). Hillsdale, NJ: Erlbaum.
- Foa, E. B., & McNally, R. J. (1986). Sensitivity to feared stimuli in obsessive-compulsives: A dichotic listening analysis. *Cognitive Therapy and Research*, 10, 477–485.
- Goldstein, A. J., & Chambless, D. L. (1978). A re-analysis of agoraphobia. *Behavior Therapy*, 9, 47–59.
- Gray, J. A. (1982). *The neuropsychology of anxiety*. Oxford: Clarendon Press.
- Gregory, W. L., Cialdini, R. B., & Carpenter, K. M. (1982). Self-relevant scenarios as mediators of likelihood estimates and compliance: Does imagining make it so? *Journal of Personality and Social Psychology*, 43, 89–99.
- Hammen, C., Marks, T., Mayol, A., & deMayo, R. (1985). Depressive self-schemata, life stress, and vulnerability to depression. *Journal of Abnormal Psychology*, 94, 308–319.
- Hibbert, G. A. (1984). Ideational components of anxiety: Their origin and content. *British Journal of Psychiatry*, 144, 618–624.
- Jannoun, L., Oppenheimer, C., & Gelder, M. (1982). A self-help treatment program for anxiety state patients. *Behaviour Therapy*, 13, 103–111.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Kendrick, M. J., Craig, K. D., Lawson, D. M., & Davidson, P. O. (1982). Cognitive and behavioral therapy for musical-performance anxiety. *Journal of Consulting and Clinical Psychology*, 50, 353–362.
- Kunst-Wilson, W. R., & Zajonc, R. B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science*, 207, 557–558.
- Lader, M. H., & Mathews, A. (1968). A physiological model of phobic anxiety and desensitization. *Behavior Research and Therapy*, 6, 411–421.
- Lang, P. J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology*, 16, 495–512.
- Lang, P. J. (1985). The cognitive psychophysiology of emotion: fear and anxiety. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders*. (pp. 130–170). Hillsdale, NJ: Erlbaum.
- Lewinsohn, P. M., Steinmetz, J. L., Larson, D. W., & Franklin, J. (1981). Depression related cognitions: antecedents or consequences? *Journal of Abnormal Psychology*, 90, 213–219.
- Lick, J. (1975). Expectancy, false galvanic skin response feedback and systematic desensitization in the modification of phobic behaviour. *Journal of Consulting and Clinical Psychology*, 43, 557–567.
- McKeon, J., Roa, B., & Mann, A. (1984). Life events and personality trait in obsessive-compulsive neurosis. *British Journal of Psychiatry*, 144, 185–189.

- MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology, 95*, 15–20.
- Marks, I. M. (1969). *Fears and phobias*. New York: Academic Press.
- Martin, I., & Levey, A. B. (1985). Conditioning, evaluations and cognitions: An axis of integration. *Behavior Research and Therapy, 23* (2), 167–175.
- Martin, M., Ward, J. C., & Clark, D. M. (1983). Neuroticism and the recall of positive and negative personality information. *Behavior Research and Therapy, 21*, 495–503.
- Mathews, A., Gelder, M. G., & Johnston, D. W. (1981). *Agoraphobia, nature and treatment*. New York: Guilford Press.
- Mathews, A. (1971). Psychophysiological approaches to the investigation of desensitisation and related procedures. *Psychological Bulletin, 76*, 73–91.
- Mathews, A., & MacLeod, C. (1985). Selective processing of threat cues in anxiety states. *Behavior Research and Therapy, 5*, 563–569.
- Mathews, A., & MacLeod, C. (1986). Discrimination of threat cues without awareness in anxiety states. *Journal of Abnormal Psychology, 95*, 1–8.
- Mayo, P. R. (1983). Personality traits and the retrieval of positive and negative memories. *Personality and Individual Differences, 4*, 465–472.
- Meichenbaum, D., Gilmore, B., & Fedoravicious, A. (1971). Group insight vs. group desensitisation in treating speech anxiety. *Journal of Consulting and Clinical Psychology, 36*, 410–421.
- Mineka, S. (1985). Animal models of anxiety-based disorders: their usefulness and limitations. In A. H. Tumas & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 199–244). Hillsdale, NJ: Erlbaum.
- Mogg, K., Mathews, A., & Weinman, J. (1987). Memory bias in clinical anxiety. *Journal of Abnormal Psychology, 96*.
- Murphy, G. E., Simons, A. D., Wetzell, R. D., & Lustman, P. J. (1984). Cognitive therapy and pharmacotherapy: Singly and together in the treatment of depression. *Archives of General Psychiatry, 41*, 33–41.
- Nunn, J. D., Stevenson, R. J., & Whalan, G. (1984). Selective memory effects in agoraphobic patients. *British Journal of Clinical Psychology, 23*, 195–201.
- Parkin, A. J., Lewinsohn, J., & Folkard, S. (1982). The influence of emotion on immediate and delayed retention. *British Journal of Psychology, 73*, 389–393.
- Parkinson, L., & Rachman, S. (1981). Intrusive thoughts: The effects of an uncontrived stress. *Advances in Behavior Research and Therapy, 3*, 111–118.
- Rachman, S., & Seligman, M. E. P. (1976). Unprepared phobias: "Be prepared." *Behavior Research and Therapy, 14*, 333–338.
- Rachman, S. (1977). The conditioning theory of fear acquisition: A critical examination. *Behavior Research and Therapy, 15*, 375–388.
- Seamon, J. G., March, R. L., & Brody, N. (1984). Critical importance of exposure duration for affective discrimination of stimuli that are not recognized. *Journal of Experimental Psychology, Learning, Memory and Cognition, 10* (3), 465–469.
- Seligman, M. E. P. (1971). Phobias and preparedness. *Behavior Therapy, 2*, 307–320.
- Stirling, N. (1979). Stroop interference: An input and output phenomenon. *Quarterly Journal of Experimental Psychology, 31*, 121–132.
- Teasdale, J. D. (1983). Negative thinking in depression: Cause, effect, or reciprocal relationship? *Advances in Behaviour Research and Therapy, 5*, 3–25.
- Von Baeyer, W. (1969). Social and political aspects of anxiety. In M. Lader (Ed.), *Studies of anxiety*. (pp. 93–98). London: World Psychiatric Association.
- Warren, R. E. (1972). Stimulus encoding and memory. *Journal of Experimental Psychology, 94*, 90–100.
- Watts, F. N., McKenna, F. T., Sharrock, R., & Trezise, L. (1986). Colour-naming of phobic-related words. *British Journal of Psychology, 77*, 97–108.
- Zajonc, R. B., & Markus, H. (1984). Affect and cognition: The hard interface. In C. Izard, J. Kagan, & R. Zajonc (Eds.), *Emotion, cognition, & behavior* (pp. 73–102). New York: Cambridge University Press.

CHAPTER 11

Cognitive Theories of Depression

M. J. Power

INTRODUCTION

What would constitute an adequate cognitive theory of depression? In order to assess extant theories and to speculate about the directions that future theories might take, this question can be broken down into at least three parts: first, the theory should draw on and stand up satisfactorily to current knowledge in cognitive science; second, it necessitates an account of depression that is clinically, empirically, and phenomenologically sound; and, third, the theory should have useful therapeutic implications. These questions in turn can be further broken down. To consider cognitive adequacy first, the theory should provide an account of at least the following: (a) the type of knowledge representation or belief system; (b) the relationship between cognition and emotion; (c) the occurrence of vulnerability and resilience; and (d) the types of cognitive processes implicated, for example, controlled versus automatic (or conscious versus unconscious) processes. Given that these preliminary cognitive modeling requirements are met, the theory then needs to account for characteristic aspects of depression such as the following: (a) the meaning of depression and its relationship to normal sadness; (b) the natural history of depression, for example, its typically time-limited course; (c) the high prevalence of depression; (d) the interpersonal and contextual aspects of depression, for example, in relation to the importance of loss and failure; and (e) specific clinical and empirical findings, for example, low self-esteem, irrational beliefs, and so on.

Although additional criteria can be added to this list, a theory that satisfies even these conditions has yet to be formulated. The aims of this chapter are, first, to consider the current theories of reformulated learned helplessness (Abramson, Seligman, & Teasdale, 1978; Peterson & Seligman, 1984), semantic network theory (Bower, 1981; Bower & Cohen, 1982; Ingram, 1984) and Beck's cognitive therapy (Beck, Rush, Shaw, & Emery, 1979) in relation to these criteria, and, second, to present some speculations that attempt to go beyond current theories.

THREE THEORIES

REFORMULATED LEARNED HELPLESSNESS

Seligman's work on learned helplessness (e.g., Seligman, 1975) and its subsequent revision (Abramson *et al.*, 1978) represents an empirically derived set of hypotheses that have been applied to depression. As every sophomore knows, learned helplessness theory originated with the finding that dogs who experienced inescapable shock later failed to escape from escapable shock, unless they were dragged across the shuttlebox by the experimenter. Seligman proposed that the experience of noncontingent reinforcement led to a state of canine helplessness analogous to depression in humans. A considerable amount of research was designed to induce helplessness in students by asking them to solve unsolvable anagrams or to escape from inescapable noise, but many of them refused to become helpless. These negative findings and the failure of the original helplessness theory to account for phenomena such as low self-esteem in depression prompted the switch from a theory that could be couched solely in behavioral terms to a cognitive theory in which helplessness resulted from the individual's explanations for perceived noncontingency rather than noncontingency *per se*. The implications of this fundamental shift (which has been glossed over in other cognitive-behavioral theories) can easily be underestimated; as Peterson (1982) has commented "the reformulation may not build on the original helplessness model so much as replace it" (p. 100).

The changes introduced by Abramson *et al.* (1978) included the proposal that individuals seek explanations for the occurrence of events and that there are critical attributional dimensions that characterize these explanations. These causal dimensions are internal-external (due to something about oneself versus something about another person or circumstances), stable-unstable (due to something that would recur for future similar events), and global-specific (whether the cause affects one or many areas of the person's life). Seligman and his colleagues predicted that depressives would be more likely to attribute bad events to internal, stable, global factors, whereas good events would be attributed to external, unstable, specific factors (Seligman, Abramson, Semmel, & von Baeyer, 1979).

The manner in which attributions lead to a risk for depression is shown in Figure 1. The occurrence of a bad event produces negative affect, which increases with the importance of the event (cf. Weiner, 1985). If the causes of the event are perceived to be uncontrollable, this should lead to passivity, the expectation of future uncontrollability, and therefore to a state of helplessness, though the extent of these emotional, motivational, and cognitive deficits depends on the associated attributions. If helplessness is perceived to be due to an internal factor, the state is termed personal helplessness. Because the individual believes that other people would not be helpless, this negative social comparison leads to lowered self-esteem. However, if the individual makes an external attribution to the effect that anybody would experience helplessness in the same circumstances the state is termed universal helplessness and no lowering of self-esteem occurs; thus, there should be a subgroup of depressives who do not show lowered self-esteem (Abramson, Garber, & Seligman, 1980). If the individual makes additional attributions to stable and global factors, the emotional,

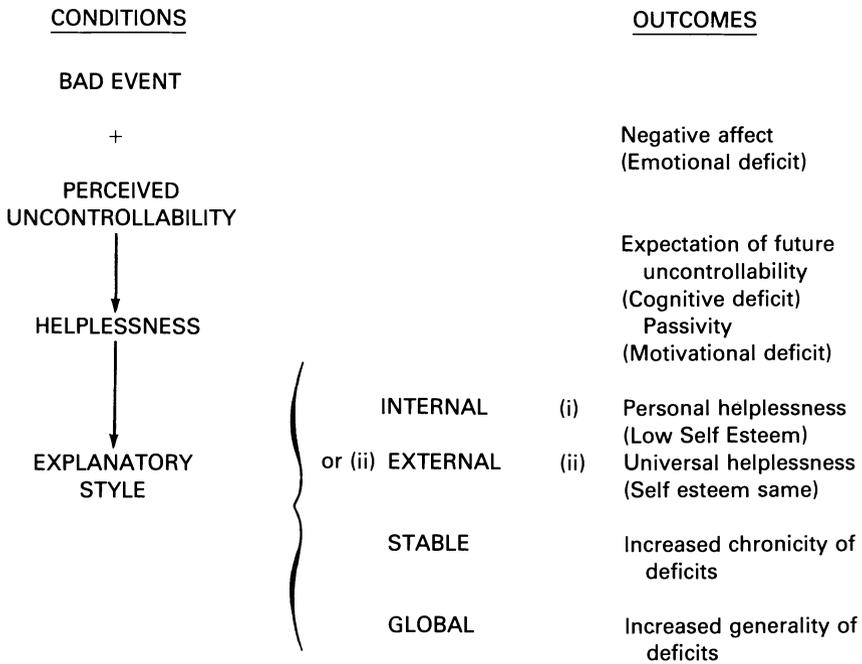


FIGURE 1. The prototypical helplessness conditions together with the resultant emotional, cognitive, and motivational deficits. The explanatory style dimensions modify the chronicity and generality of these deficits.

motivational, and cognitive deficits become more chronic and general in their effects.

In order to examine the implications of the theory more fully two further sets of conditions will be considered. The first of these is presented in Figure 2. In this case a good event occurs, therefore there is no negative affect (e.g., Abramson *et al.*, 1980). Nevertheless, the perception of the event to be uncontrollable should lead to the cognitive and motivational deficits; that is, helplessness without negative affect. If the individual makes an internal attribution for this event, this should lead to an increase in self-esteem (cf. Weiner, 1985). Although this outcome may seem counterintuitive—a nondepressed helplessness with increased self-esteem—it may occur with cases of “golden girls” and “golden boys” (Seligman, 1975) and in Bruch’s (1978) “golden cage” of anorexia nervosa. One of the reasons for the relative rarity of this type of helplessness may be that, as the illusion-of-control studies show, people normally perceive good events as controllable even when the events are objectively uncontrollable (e.g., Alloy & Abramson, 1979; Alloy & Tabachnik, 1984). Nevertheless, the repeated occurrence of noncontingent good events can result in an individual with much apparent privilege, but little or no sense of control; one of the problems is that bad events will eventually occur.

A further example of the application of the reformulated theory is shown in Figure 3. In this example an individual perceives a bad event to be controllable; thus, there is negative affect, but no motivational deficit (passivity) or cognitive deficit

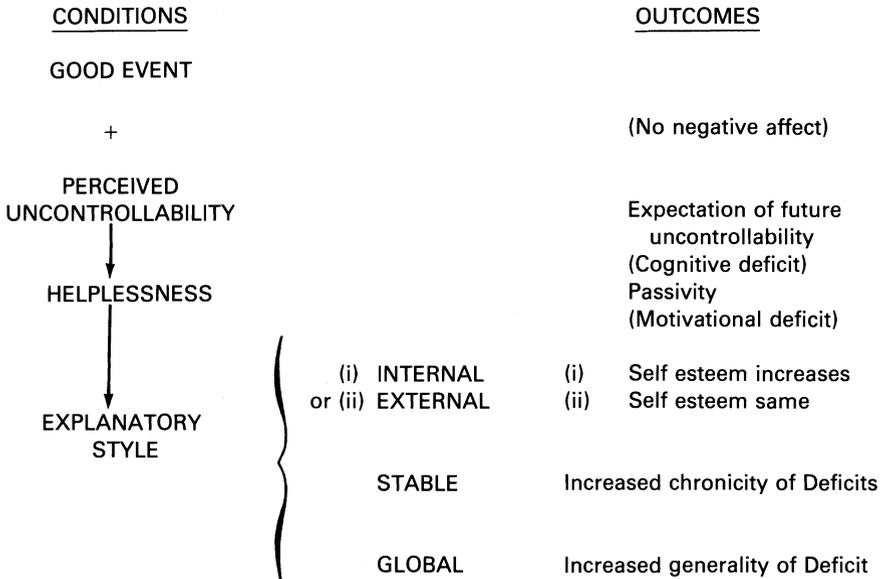


FIGURE 2. The diagram shows a good event perceived as uncontrollable, which should lead to helplessness without negative affect and an increase in self-esteem if the causes of the event are seen to be internal.

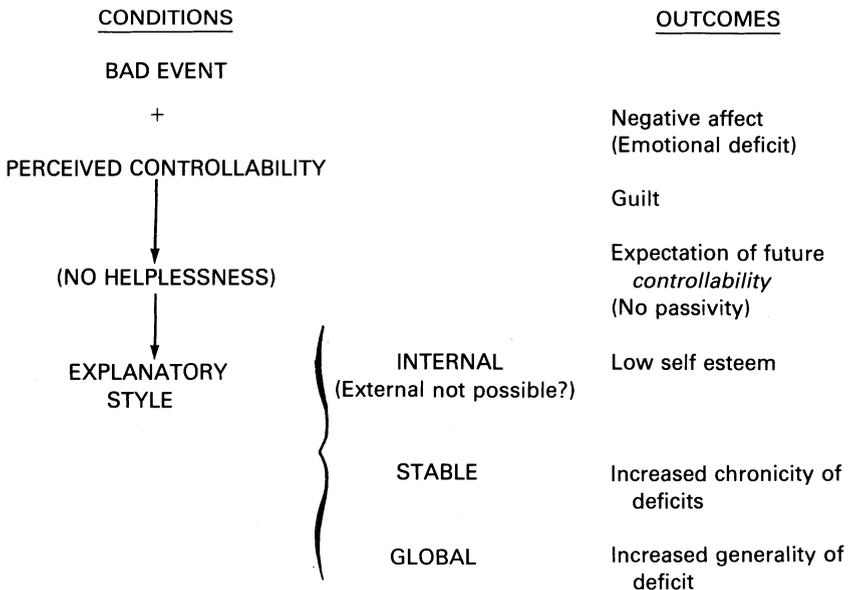


FIGURE 3. The perception of a bad event as controllable should lead to an emotional deficit that includes guilt, but no cognitive or motivational deficits.

(expectations of future uncontrollability). If the individual makes an internal attribution, this should lead to lowered self-esteem. However, with the model it does not seem possible to make an external attribution to an event that is perceived to be controllable, because by definition the individual must perceive that controllable outcomes are contingent on his or her own responses. Therefore, the emphasis on perceived controllability within the reformulated theory means that the attributional dimensions are not logically separate or orthogonal to the controllability dimension. In fact, it must be noted that the Attributional Style Questionnaire (Peterson *et al.*, 1982), which was designed as a measure of attributional style, fails to include a measure of controllability, even though it is measured by other attributional theorists (e.g., Weiner, 1985). These considerations suggest that there are logical inconsistencies within the theory.

A number of main predictions from the reformulated theory have received considerable empirical testing. Peterson & Seligman (1984) summarized much of this work and concluded that the evidence is overwhelmingly in favor of the reformulation, whereas other researchers have argued for the more cautious conclusion that although there is some correlation between the attributional dimensions (or "explanatory style" dimensions as Peterson and Seligman rename them) and depression there is no good evidence that the predicted style is causal in the onset of depression (e.g., Brewin, 1985; Coyne & Gotlib, 1983). Brewin (1985) and Power (in press) have further agreed that other dimensions, such as the perceived moral aspects of events and the consequences rather than the causes of events, may be more important in the development of depression, and Weiner (1985) has questioned the need for the global dimension.

In terms of the other criteria the reformulation does well on some points but badly on others; the theory meets the requirements for an adequate account of depression better than it meets the requirements for an adequate cognitive theory. An interesting set of emotional, motivational, cognitive, and self-esteem deficits are presented that can appear in a number of combinations, though the emphasis on helplessness depression may have distracted from considerations of how helplessness relates to other conditions: helplessness does not necessitate depression, nor does depression necessitate helplessness. However, whether these proposals prove to be empirically and clinically sound in their account of the phenomena of depression has yet to be decided, but the narrow focus on the causal aspects of events seems both theoretically and empirically unjustified.

A further important weakness with the reformulation is that it fails to include an adequate cognitive theory (Power & Champion, 1986). There is no account of the form of internal representation on which the theory is based; for example, the theory could incorporate Bower-type semantic networks, Beck-type schemata, or any other such cognitive structures. Nor does the theory consider the cognitive processes that underlie the making of attributions; for example, unconscious processes could lead to one set of attributions that were consciously altered so that individuals presented themselves in a favorable or (in the case of depressives) unfavorable light.

A final and most important requirement of a theory of depression is that it should have implications for therapeutic practice. The following chapter by Williams considers cognitive therapies in detail, but it must be commented that some of the

therapeutic procedures derived from the reformulation may be problematic, in particular the procedures of personal-control training and attribution retraining (e.g., Seligman, 1981). The problems are that under certain conditions depressives are more realistic (e.g., Layne, 1983; Lewinsohn, Mischel, Chaplin, & Barton, 1980; Power & Champion, 1986), show less self-deception (Roth & Ingram, 1985), are more accurate in their perception of contingency and noncontingency (e.g., Alloy & Abramson, 1979), sometimes set more realistic expectations (e.g., Golin, Terrell & Johnson, 1977), are more accurate in their recall of punishment (e.g., Nelson & Craighead, 1977), and experience more negative events in the first place (e.g., Krantz, 1985). These results suggest that a depressed patient's attributions and beliefs are sometimes more accurate and realistic than the therapist's (especially if the therapist is not depressed); therefore, a therapist who exhorts a patient to adopt a less realistic set of beliefs may be unlikely to succeed.

SEMANTIC NETWORK THEORIES

A general theory of the relationship between cognition and emotion has been presented by Bower and his colleagues (Bower, 1981; Bower & Cohen, 1982). A similar account has been presented by Teasdale (e.g., 1983), and Ingram (1984) has further adapted Bower's general theory to apply specifically to depression. The general approach is that concepts, events, and emotions are represented as nodes in a network of interconnections; a small portion of such a network is shown in Figure 4. There are a number of important features. First, the depression emotion node or "DEMON" (Power & Champion, 1986) has a wide variety of types of linkages. There are links to expressive behaviors and autonomic patterns, some of which Bower & Cohen (1982) consider to be innate. The DEMON also has links to other emotions: Figure 4 portrays an inhibitory linkage between depression and happiness, which captures the fact that certain emotions oppose each other in their effects. The diagram also shows so-called evoking appraisals or perceptual inputs that activate the DEMON. Other

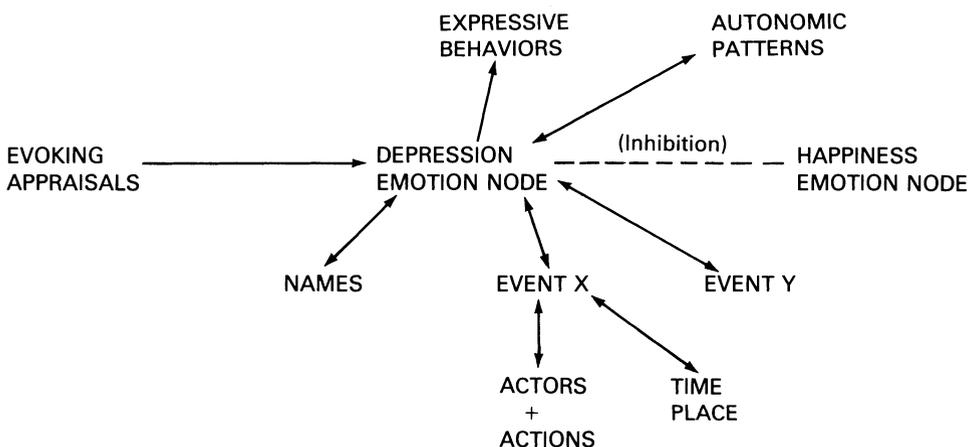


FIGURE 4. A fragment of semantic network that includes the depression emotion node (adapted from Bower & Cohen, 1982).

linkages include verbal labels that are used to name the emotion, and sets of associated events that have become linked through experience with the DEMON (e.g., see Teasdale, 1983). These events are represented in the form of propositional structures, the details of which are not shown in the diagram but that can be seen in Bower & Cohen (1982). A second feature of the network is that activation of one node leads to the spread of activation to associated nodes; thus, the evocation of the DEMON by an appropriate situation, such as failing at an important task, leads to a spread of activation to events that have become associated with the DEMON. Those events reaching a high enough activation level will enter consciousness. An analogy that Bower has made for this spread of activation is an electrical circuit in which the application of voltage at one terminal spreads to adjoined terminals along interconnecting wires of different resistances. However, the inclusion of inhibitory connections, and critical thresholds for certain connections (Bower & Cohen, 1982) makes this simple analogy inadequate.

The network theory presented so far provides an account of automatic processes by which cognition and emotion may be linked, but it does not address the processes that Posner & Snyder (1975) labeled as controlled or conscious. Bower & Cohen (1982) and Gilligan & Bower (1984) have presented several additional features for their network model that incorporate this controlled-automatic distinction. The main additions are a blackboard control structure and sets of interpretation rules. The blackboard control structure is equivalent to the more commonly used concept of working memory (e.g., Baddeley & Hitch, 1974; Hitch, 1980; Power, 1985); therefore the latter term will be used in preference. Working memory provides a workspace that can be flexibly distributed between temporary storage and processing requirements; thus, a preliminary representation of a relevant scene or event can be held in working memory and this representation can be transformed by various automatic and controlled processes (see Figure 5). According to Bower and his colleagues, sets of cognitive-interpretation, emotional-interpretation, and emotional-interaction rules

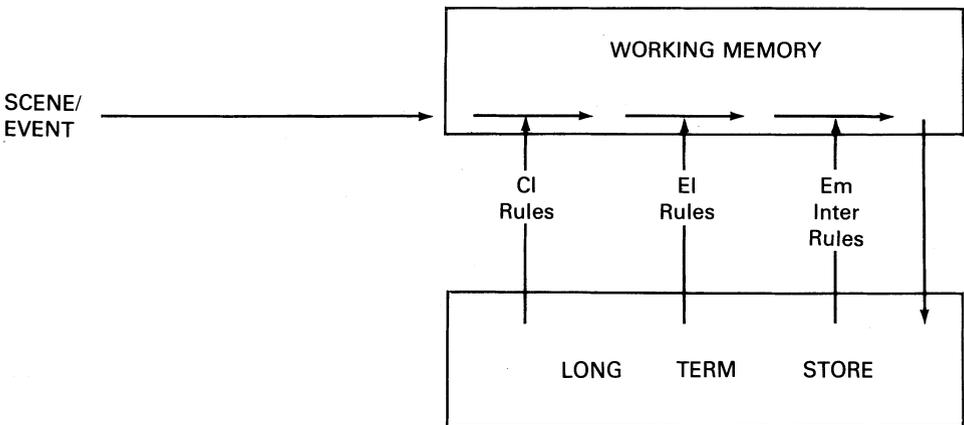


FIGURE 5. Schematic operation of the Bower & Cohen (1982) model in which cognitive-interpretation (CI), emotion-interpretation (EI), and emotion-interaction (Em. Inter.) rules operate on representations of a scene or event in working memory.

operate on the initial representation and the products of various intermediate stages (see Figure 5). These rules are held in Long Term Store (LTS) in a conditional format "If . . . , Then . . . ," for example "If there is a potential for harm, then increase fear" (Bower & Cohen, 1982). The rules provide for a much richer interface between the input and the semantic network.

Despite the revisions in the Bower & Cohen (1982) network theory, there are a number of major problems. First, a semantic network is not a theory in itself, but merely a format in which a theory can be expressed; thus, the representation of a sentence or an event within a Bower semantic network still requires a semantic interpretation by means of procedures that can assign truth conditions to the propositions (Johnson-Laird, 1983). A second and related problem is that the units of representation within the network are inappropriately small for capturing the structure of certain events, actions, and situations for which molar units of representation, such as scripts, schemata, or mental models are more useful; network theories were originally designed to represent the relationship between individual words whereas other domains of knowledge are more usefully organized in larger units. Other problems with the cognitive adequacy of the theory have been considered by Power & Champion (1986)

The least impressive aspects of Bower's theory have been the account provided of depressive phenomena and the theory's implications for therapeutic practice. Unlike reformulated helplessness or Beck's cognitive therapy, the theory has nothing specific to say about depression; to a certain degree it is merely a redescription of the observation that certain emotions and thoughts tend to be found together. This theoretical emptiness could be overcome by importing content from other theories of depression, but the cognitive basis of the theory needs to be reconsidered first.*

BECK'S COGNITIVE THERAPY

Beck and his colleagues have developed a cognitive theory of depression and other disorders (Beck, 1976; Beck *et al.*, 1979), which, in contrast to the laboratory based approaches of Seligman and Bower, has been primarily based on clinical data. It is not surprising therefore that the two major requirements that the approach account for the phenomena of depression, and that it should have major therapeutic implications are met better by Beck's approach than by either Seligman's or Bower's. Instead, the main problems to be considered will focus on its adequacy as a cognitive theory, though some of these issues will have implications for therapeutic strategy.

Beck's theory can be considered in two parts. First, there is the structural part of the theory, in which a hierarchically ordered set of schemata are the units of representation in which information about the self, the future, and the world is represented. Second, there is a set of information-processing strategies that are considered to lead to typical depressive distortions or logical errors in thinking. Although these two parts are integrated in the theory, it is possible to accept one without the

*It has been informally reported to us that some patients have found the presentation of an outline of Bower's theory a useful educational adjunct in cognitive therapy. However, this use is not relevant to the question of the theory's accuracy.

other. The choice of schemata as units of representation together with a “thinking is rational and logical” assumption may give rise to internal contradictions within the theory: the schema concept was originally imported into psychology because memory and perception are inherently distorted (Bartlett, 1932). A surprising conclusion therefore may be that cognitive therapy would be a more internally consistent theory if it were based on semantic networks rather than on schemata; whereas schemata are inherently distorting, networks are not.

The structural part of the theory proposes that the depressive has a characteristic cognitive triad of beliefs in which the self is seen as negative, the future consists of unremitting hardship and failure, and the world is seen to be the terrible place that it is. This set of underlying assumptions about the self and the world is considered to result from problems in significant early relationships (cf. Bowlby, 1980; Brown & Harris, 1978) that lead to a continuing psychological vulnerability. The subsequent occurrence of related events or situations may then give rise to an increase in the negative processing of information and to depression (see Figure 6). In between episodes of depression these dysfunctional schemata are considered to be latent; as Sacco & Beck (1985) state, “The depressogenic cognitive schemas will remain latent until activated by stressors (precipitating factors) to which the individual is sensitized” (p. 5).

Beck (1983) has further proposed that vulnerability may be divided into two subtypes. First, there are the autonomous individuals who typically show high confidence and good self-esteem, high self standards, independence, and goal orientation and who tend to distance themselves from others when not depressed. Second, there are dependent or sociotropic individuals who need others for security and overcoming fears of abandonment, and who constantly seek closeness. Although these proposals are based primarily on clinical observations, similar subtypes of depression have been considered by Arieti and Bemporad (1978) and by Blatt and his colleagues (e.g., Blatt, D’Affliti, & Quinlan, 1976).

A number of problems with Beck’s theory will now be outlined. First, the concept of schemata in the sense used within the theory is extremely vague (cf. Ingram, 1984), and, indeed, the term is frequently considered to be synonymous with beliefs, attitudes, and assumptions. This vagueness, although it may satisfactorily express the clinical aspects of depressive vulnerability, is unnecessary given the relatively detailed consideration that the concept has received within cognitive science (e.g., Mandler, 1984; Rumelhart & Norman, 1985).

Second, there are alternatives to Beck’s proposed hierarchical organization of schemata and to the proposed latency of schemata between episodes of depression. For example, while at work an individual’s primary belief might be in the need to compete successfully at all costs, and a subsidiary belief might be in the need to be loved by other people, whereas in an intimate relationship the reverse might hold

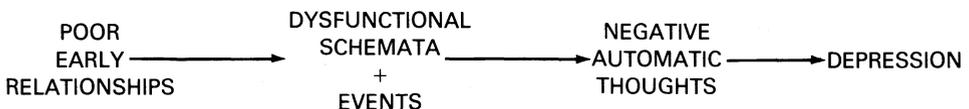


FIGURE 6. The pathway to depression within Beck’s cognitive theory.

and the predominant belief might now become the need to be loved. This alternative heterarchical organization therefore includes the properties of a hierarchy while adding a considerable flexibility that can be responsive to events and to situations. Similarly, it seems unlikely that critical schemata concerned with issues of loss and failure are latent between episodes of depression, given that they are of such importance to the individual. The proposal that they are latent seems to be *ad hoc* and based on the failures to find the self-report of dysfunctional attitudes in recovered depressives (e.g., Hamilton & Abramson, 1983; Wilkinson & Blackburn, 1981). An alternative *ad hoc* account is that active processing does occur but the outcomes of this processing are inhibited by either controlled or automatic processes (Power & Champion, 1986).

Third, Beck's assumption that normal thinking is more rational, logical, and realistic than depressive thinking is challengeable on a number of grounds. In the discussion of Seligman's reformulated helplessness it was pointed out that depressed patients are sometimes more accurate and realistic than their normal counterparts. These results, together with investigations of normal reasoning (e.g., Johnson-Laird, 1983) suggest that the question of whether depressives are more rational than normals or whether normals are more rational than depressives is a misguided one. As Power & Champion (1986) suggest, depressives may be more accurate with negative information that is correct because the conclusions that they draw are congenial with their models of a negative view of the self, whereas normals are more accurate with positive information that is correct because these conclusions are congenial with their positive models of the self. Related examples of how beliefs can influence reasoning may be found in Oakhill and Johnson-Laird (1985). One of the implications of this proposal is that, as pointed out earlier, a cognitive therapy strategy in which negative conclusions are challenged may sometimes be more irrational and more illogical than the reasoning by which the negative conclusions are reached. The difficulty in therapy is that depressives are painfully right about some of the negative aspects of their situation, but painfully wrong about many of the positive aspects. In contrast, normals either in real life or in therapy are right about many of the positive aspects, but annoyingly unaware of many of the negative.

SOME SPECULATIONS

Now that everyone else has been pushed off the wall, the only problem left is to put Humpty Dumpty back together again. In attempting to succeed where so many horses and men have failed, the starting point will be to build as secure a cognitive foundation for the theory as possible. Specific aspects of depression will then be considered, and, finally, there will be a brief discussion of the course of depression.

A PUTATIVE COGNITIVE BASIS

In this section three particular aspects of current cognitive science will be highlighted which, it will be argued, can overcome some of the limitations of the

cognitive theories considered earlier. The three aspects are as follows. First, the structural part of the theory will take a mental model approach (Johnson-Laird, 1983), which provides a higher level of representation than units such as propositions or schemata. Second, an important distinction needs to be made between controlled and automatic, or conscious and unconscious processes (e.g., Posner & Snyder, 1975). And, third, a cognitive theory of emotions will be adopted in which emotions are seen to arise at different junctures in plans and goals (e.g., Miller, Galanter, & Pribram, 1960; Oatley & Johnson-Laird, 1985). Each of these three proposals will now be discussed in more detail.

Mental Models. The theory of mental models developed by Johnson-Laird (1983) has primarily focused on the areas of syllogistic reasoning and discourse comprehension, but it can be readily extended to form a general theory of mind. There are three parts to the theory: a propositional level of representation, a higher level of mental models, and a set of procedures that relate propositions and models to each other. The first part of this section will be devoted a relatively technical discussion of the theory of mental models; in the latter half of the section and in subsequent sections consideration will be given to the application of the theory to depression.

The propositional level of representation is an intermediate level between mental models and language; models can take the form of images where there is a perceptual input, and some people can construct images from verbal description via this intermediate propositional level. Although there are many uses of the term *proposition* it will be taken here to imply the smallest semantic unit for which a truth value can be stated; thus, the word *god* is in itself neither true nor false, but the sentence "god is dead" can be stated to be true or false in different possible worlds. In this philosophical usage therefore a propositional representation is the mental representation of an entity that can be expressed by a sentence and it is either true or false of the state of affairs that it describes. One example of the use of propositional representations is in Bower's semantic network described earlier in which sentences are represented in the form of propositions through the addition of links and pointers between the appropriate nodes in the network.

There is good empirical evidence for the psychological reality of a propositional level of representation (e.g., Johnson-Laird, 1983; Power, 1986) but there is also good evidence that the cognitive system operates with a higher level of representation that has been referred to in various ways by other authors, but that Johnson-Laird has termed a mental model. Mental models are typically multidimensional and consist of sets of tokens that represent entities (e.g., dog, justice, happiness), a set of properties or attributes of these tokens (e.g., red, tall, cowardly) and a set of relations between the tokens. One way of envisaging such structures is like a chemical model of a molecule (Power & Champion, 1986), though as Johnson-Laird observes the appearance of models "in the head" may never be known. The point is that, following Alan Turing's criteria, the cognitive system behaves as if it operates with mental models and propositional representations.

The third component is a set of procedures. These procedures are necessary to construct models out of propositions, and, in reverse, to produce propositions from models. These procedures may initiate the construction of a new model or the verification or modification of an old model. In the construction of a model to represent

a state of affairs, arbitrary inferences may need to be made that may subsequently prove to be incorrect; that is, much information is ambiguous and indeterminate and certain interpretations and arbitrary inferences may need to be made to construct a mental model. For example, in the garden path sentence "The boat floated down the river sank," the initial model of a boat floating by its own accord needs to be revised by a procedure that introduces a relative clause into the sentence so that it becomes "The boat which (was) floated down the river sank." Even now a further procedure may be needed if the passive object relative clause *which was* is incorporated into the new model in order to introduce an unspecified agent who initiated the floating. Procedures that evaluate models and check whether certain alterations make the models consistent with new information are recursive in that a model needs to be constructed that includes a representation of itself. These general recursive procedures provide computational power: a deliberate comparison in that, as Turing proposed, any effective procedure can be computed. The role of the computer in cognitive science is therefore not as many mistakenly think that of current metaphor for mind, but instead any effective theory of the mind should be computable.

Two main varieties of models are considered in the theory; namely, physical models and conceptual models. Physical models correspond to spatial, temporal, and dynamic aspects of the physical world, though in addition to representations of physical entities and relationships they also incorporate an abstract concept of causation. Conceptual models are more relevant for theories of depression. They capture abstract relations such as negation, conjunction, and disjunction; examples of conceptual models include assertions about individuals, abstract relations between individuals, quantification, and group membership.

To give an example of a mental model, the assertion 'Cognitive theories of depression are wrong' could be represented by many types of models that include the following two:

<i>Model 1</i>	CTD = Wrong	<i>Model 2</i>	CTD = Wrong
	CTD = Wrong		CTD = Wrong
	CTD = Wrong		CTD = Wrong
			(CTD) (Wrong)

In these models CTD (Cognitive Theory of Depression) is a token, "Wrong" is a property of a token, and "=" represents the relation between them. Model 1 assumes that there are three cognitive theories, each of which is wrong, whereas Model 2 includes an imaginary world in which there is a cognitive theory of depression that is not wrong (indicated by the unattached CTD); it also allows for the possibility that things other than cognitive theories can be wrong. Both models represent the initial assertion, but because this assertion is indeterminate different arbitrary inferences have had to be made before a specific model could be constructed.

There are a number of properties of mental models that make them appropriate for a theory of depression. First, mental models provide a higher level of representation than either schemata or the propositions incorporated in a semantic network. Therefore, this approach could take a semantic network as its starting point and it could

incorporate schema-type information into the construction of mental models. This higher level of organization seems more appropriate to the general themes that concern the self: the world, goals, relationships, loss, and failure, which are central in depression. Second, the focus of mental models is on their usefulness rather than on their truthfulness or accuracy. This property does not require that normal thinking is rational and logical, nor that depressive thinking is irrational and illogical. Instead, it predicts that under certain conditions both normal and depressive thinking will appear rational and logical; as outlined earlier, depressed individuals with negative views of themselves may draw rational and logical conclusions about negative information, because the conclusions are congenial to their knowledge of themselves and the world. However, normal individuals are more likely to reject negative information about themselves and seek further explanations for these noncongenial negative conclusions (cf. Schwarz & Clore, 1983). Third, mental models may be embedded within each other and meet the earlier stated requirement of heterarchical organization. These properties enable one model to appear superordinate to another under some conditions, but subordinate to that same model under other conditions. One application may be that the model of the self in depressives may switch between being predominantly negative during an episode of depression, but be predominantly positive between episodes of depression (see later).

Conscious and Unconscious Processes. A crucial distinction needs to be made in any cognitive theory between conscious (controlled) and unconscious (automatic) processes (e.g., Bowers & Meichenbaum, 1984; Posner & Snyder, 1975; Power, 1985; Schneider & Shiffrin, 1977). The operation of automatic processes is typically modular, parallel, and fast, whereas controlled processes are usually sequential and slow. However, controlled processes do not act independently of automatic ones; they may call on automatic processes for their operation (e.g., in fine motor control), they may take control of processes that are normally automatic (e.g., in the conscious control of breathing or walking) (cf. Shiffrin & Schneider, 1984), they may alter or reject the outcomes of automatic processes, or they may facilitate or inhibit such processes (cf. Neely, 1977; Posner & Snyder, 1975). The degree of penetrability of automatic processes by conscious ones may vary across modalities, such that some processes are impenetrable, though it is possible to construct a mental model of their operation (cf. Johnson-Laird, 1983), whereas other processes may be much more transparent to conscious control and interference (cf. Dixon, 1981). This proposal is a modification of the Nisbett and Wilson (1977) argument that mental processes are completely impenetrable and that, therefore, self-reports are unrepresentative or inaccurate sets of data: the conclusion is that it all depends.

This distinction between conscious and unconscious processes has widespread implications for a cognitive theory of depression. For example, facilitatory and inhibitory effects are necessary within any theory of emotion to account for different types of interactions between emotions (e.g., Bower & Cohen, 1982). Similarly, as argued earlier, between episodes of depression the depressive may be able to inhibit the entry into awareness of the automatic processing of negative material, whereas during depression conscious processes appear to facilitate this negative processing.

The implications of the distinction are also evident in therapy. Cognitive-behavioral strategies that require the patient to make positive self-statements may

make little impact on behaviour that is controlled by automatic processes. In contrast, patients may often proclaim toward the end of therapy that they have failed to make any gains, because they are unaware of the progress that is apparent to the therapist and to significant others. A more chronic user of this denial is the so-called characterological depressive (e.g., Akiskal, Bitar, Puzantian, Rosenthal & Walker, 1978) who presents a negative plaintive set whatever the weather, but who, in fact, like the central character in Robert Musil's *The Man Without Qualities*, often possesses qualities in abundance.

In summary, we would disagree with Sacco & Beck's (1985) statement that "the concept of unconscious processes is largely irrelevant to cognitive therapy" (p. 5) and state that the importance of unconscious processes cannot be overemphasized.

Cognitive Approaches to Emotion. Complaints about the lack of emotion in cognitive theories are now abundant. However, the textbook distinction between cognition and emotion seems false in that human emotions are not simply a primitive throwback to a lower evolutionary state, but they require the complete cognitive apparatus for their full expression (e.g., Lyons, 1980; Mandler, 1984). The proposal is that although there may be relatively undifferentiated emotional states of a positive or negative tone (e.g., Weiner, 1985), an emotion like depression requires both a general negative state plus an additional cognitive interpretation of that state.

There are a number of cognitive theories of emotion that have begun to take these considerations into account. One such example is the theory being developed by Oatley and Johnson-Laird (1985), which will be considered here to demonstrate how an emotion such as depression requires the involvement of high-level cognitive processes. Oatley and Johnson-Laird have argued that in addition to Mandler's (1984) proposal that emotions act as interrupt signals to current goals and plans, emotions also serve to *maintain* the system in a specific state; thus, in a system of multiple goals, they lead to an exchange of priorities. For example, the loss of a job may lead to sadness, to some goals being set aside, and, if circumstances permit, a new priority goal may become the search for another job. In more general terms, emotions are considered by Oatley and Johnson-Laird to be the consequence of junctures in plans: sadness results from a major loss or failure; anxiety results when self-preservation is threatened; anger occurs when an active plan is frustrated; and happiness occurs when subgoals are being achieved. For sadness to turn into the more complex state of depression, the individual's interpretation of both the meaning of the failure or loss and the subsequent basic emotions is important. Oatley & Bolton (1985) have further argued that if the failure or loss is of something or someone through which the individual defines his or her self-worth and if there is a lack of alternative sources of self-definition then depression is likely to occur. An essential part of depression therefore is not simply the loss of a goal and the resultant sadness, but the lost goal must be closely related to the individual's self-concept.

This cognitive theory of emotion will undoubtedly be modified with future developments. One of the problems is that although junctures in plans and goals may be a major source of emotions, they are not the only source (e.g., Weiner, 1985); for example, emotion can be experienced through the recollection of a pleasant or unpleasant event, without any disruption to current goals and plans; natural and

induced biochemical changes in the brain may also give rise to basic euphoric or dysphoric states, though the exact emotion experienced will further depend on the cognitive interpretation of the initial state and its purported causes. Nevertheless, the fact that the most common category of depression is event related requires a theory along the lines that Oatley and Johnson-Laird are developing.

VULNERABILITY TO DEPRESSION

In the previous section an outline was given of the general cognitive structures and processes that must underlie a cognitive theory of depression; the purpose of this section is to provide the additions that are necessary to make this a theory that is specific to depression. A number of points have been mentioned in preceding sections and these will be summarized under three headings; namely, it will be proposed that the depression-prone individual has a narrow range of goals and options, an ambivalent attitude toward the self, and high discrepancies between actual and ideal achievements and relationships. These vulnerability factors are proposed to be characteristic of an individual prior to the experience of an episode of depression; it will be argued subsequently that the experience of depression may lead to either positive or negative changes that can override some or all of these factors. A more detailed discussion can be found in Champion & Power (1987).

A Narrow Range of Goals and Options. This factor is based on the interplay of individual and social influences; it represents the fact that many of our important goals and plans are social in nature (e.g., Champion, 1985; Oatley & Bolton, 1985; Oatley & Johnson-Laird, 1985) and that even goals that are not obviously social either require a social context in which they can be achieved or they arise as a consequence of earlier social relationships.

The main proposal is that the depressive tends to have a restricted or narrow range of goals and plans and that these typically focus on either interpersonal or achievement themes (Arieti & Bemporad, 1978; Beck, 1983; Blatt *et al.*, 1976). To alter Arieti & Bemporad's (1978) terminology slightly, depressives tend to have either an interpersonal or achievement-oriented dominant goal; the characteristics of each one closely resemble the earlier discussion of Beck's sociotropic and autonomous types. In addition, it is proposed that depressives perceive that they have a narrow range of social resources and options by which to attain their goals (Freden, 1982). This restriction in social options can arise from an actual lack of options, from the inability of the individual to use resources that do in fact exist, from the lack of perception of resources, or most likely, from a combination of all of these. For example, one woman might become depressed following the loss of her spouse because of a belief that he was irreplaceable, whereas another woman might only become depressed if all of her sources of support were actually removed.

The main evidence in favor of this proposal has arisen from the clinical insights of psychotherapists such as Arieti and Bemporad (1978). However, some supportive empirical evidence has recently been reported by Hammen, Marks, Mayol, and DeMayo (1985) in a longitudinal study of mild depression in students. They found that the students could be grouped into dependent and self-critical types, plus a

smaller mixed group who showed characteristics of both. Over a 2-month follow-up dependent students were more likely to become depressed following interpersonal events, whereas for self-critical students achievement-related events were more important. Together with the clinical data, this finding suggests that life-events research in future must take account of distinctions in which the type of event is related to its importance for the individual's goals and plans.

An Ambivalent Self. In the earlier discussion of heterarchical and hierarchical organization, it was argued that one of the advantages of a heterarchical organization was that it could provide a better description of how depressives can often present as positive, confident, and with high self-esteem when not depressed, but see themselves as the opposite when depressed (see also the earlier discussion of Beck's autonomous type). Current success in the pursuit of the dominant goal provides a source of self-worth and negative aspects of the model of the self may be partially or totally inhibited. However, under threat to the dominant goal, its protective aspects are absent and the negative aspects of the self may become more difficult to inhibit: a process that reaches its height with the irrevocable loss of the person or ambition that was the dominant goal's focus. At this stage, the negative aspects dominate the model of the self and lead to the various biases in memory, perception, and reasoning (e.g., Beck *et al.*, 1979).

There are now several studies that have demonstrated that depressives report more negative attitudes about themselves and the world when they are depressed, yet they may be indistinguishable from normal controls when recovered (e.g., Hamilton & Abramson, 1983; Wilkinson & Blackburn, 1981). These findings concur with clinical impressions and support the proposal for an ambivalent model of the self in depressives. However, a more interesting proposal that has yet to be tested is whether recovered depressives show automatic negative processing effects even though consciously they may report only positive effects, and, perhaps counterintuitively, whether during episodes of depression depressives show automatic positive processing effects even though they report primarily negative attitudes.

High Discrepancies between the Actual and the Ideal. It was pointed out in the discussion of goals and options that restrictions can arise from the individual's perceptions, from an actual lack of resources in the environment, or, most commonly, from an interaction between the two; a similar argument holds for the discrepancies that depressives perceive between their actual relationships or ambitions and those they would ideally like to have. That is, there are many perfectionistic depressives whose standards are so high that no relationship or achievement is ever good enough. Conversely, there are others who seem to want nothing more than the rest of us, but who are dogged by ill luck from the start.

Two areas of research in which high discrepancies between the actual and ideal are evident (other than the literature on self versus ideal self) are social support and achievement. In the area of social support Henderson, Byrne, and Duncan-Jones (1981) reported that in a nonsymptomatic community sample a significant factor in the development of neurotic problems was a lower perceived adequacy of significant relationships. Similarly, Lewinsohn, Larson, and Munoz (1982), in their treatment outcome study of depression, found that depressives were more likely to subscribe to beliefs about personal failure and the impossibility of achieving personal happiness.

THE PROCESS OF DEPRESSION

A person may have all of the vulnerability factors listed earlier, but need never become depressed; the argument is merely that he or she is at greater risk. The process through which a vulnerable individual becomes depressed can only be speculated about at present. Undoubtedly, life events and chronic difficulties play a major role in this process (Brown & Harris, 1978), but it is less clear which factors lead from normal sadness or everyday misery into depression; which factors predict the severity of depression; and which factors lead to a good prognosis for some individuals and to a chronic course for others. Each of these problems will be briefly considered.

Research has shown that depressed individuals experience significantly more life events that involve loss or the threat of loss (e.g., Brown & Harris, 1978; Johnson & Sarason, 1979; Paykel, 1979). However, in spite of the overwhelming evidence for the role of life events they only account for about 10% of the variance in the severity of distress (Cochrane & Sobol, 1980; Lloyd, 1980). In order to increase the predictive power of life events, events must be related to the individual's goals and plans; as discussed earlier, events that disrupt a dominant goal will be more likely to lead to depression especially if the individual has few other options that are valued or has few resources with which to construct new goals.

The occurrence of an event that disrupts a dominant goal is predicted to lead to a disinhibition of the negative aspects of the ambivalent self in a depression-prone individual; that is, the protective function of the dominant goal no longer serves as a source of self-worth. The sadness that most people feel at the loss of someone or something valued sets in motion additional feelings of a loss of control and a heightened self-awareness (e.g., Ingram & Smith, 1984; Mechanic, 1986; Smith & Greenberg, 1981). A wide range of activities that were related to the dominant goal are now either dropped or carried on but with a sense of futility, whereas previously inhibited thoughts, feelings, and activities increase in frequency. In the most severe cases, the dominant goal may be replaced by an autonomous delusional model: a Pyrrhic victory over the flood of negative thoughts and feelings.

The experience of an episode of depression may lead to significant changes in the individual and in his or her social resources. For example, in the search for explanations for their symptoms individuals may develop predominantly physical, psychological, or social models (cf. Caine, Wijesinghe, & Winter, 1981; Jones, 1983; Mechanic, 1978, 1986), but if personal explanations do not match those of significant others, then a poorer prognosis would be predicted. After recovery, the experience of such differences, together with the appraisal of the episode itself and the nature of some of the treatment regimes may all combine to shift the model of the self in a more negative direction. For those depressives who remain predominantly negative in their self-image after recovery, the role of psychosocial factors such as life events and social support should become of decreasing importance in the onset of future episodes (Billings & Moos, 1984). Greater risk of relapse will also occur if the episode leads to further losses of close relationships and significant roles; for example, weaknesses in a relationship may become more apparent under considerable stress, or individuals may lose their jobs because they have been depressed. Such disadvantage may accumulate with repeated episodes.

CONCLUSIONS

The ideas of Seligman, Bower, and Beck have made significant contributions to theory and therapy in depression. However, each theory has weaknesses either in terms of its cognitive adequacy or in terms of its related therapeutic practices. The combination of a mental models approach (Johnson-Laird, 1983) that subsumes propositional and schema information, a distinction between conscious and unconscious processes, and a cognitive theory of emotion can overcome many of the limitations of these theories and provide a basis on which to develop a more adequate approach. The addition of depression-specific content can then be divided into the problem of vulnerability to depression and the problem of depressive process. Vulnerability can be seen to be based on a narrow range of valued or dominant goals, little or no valued social resources, such as close relationships, a highly ambivalent model of the self, and high discrepancies between actual and ideal ambitions and relationships. Such an individual is particularly at risk for depression if an event occurs that threatens or disrupts the dominant goal; the loss of the goal can lead to a decrease in a range of activities directly and indirectly related to that goal, to an increase in self-awareness, and to a consequent loss of inhibition over negative thoughts and feelings. Together, these factors can shift the vulnerable individual from everyday misery into a state of depression, which in turn may lead to long-term changes in the individual's personality.

The theory presented here is extremely speculative, but it must be noted that Lewinsohn, Hoberman, Teri, and Hautzinger (1985) have recently modified Lewinsohn's behavioral theory of depression and have independently presented a cognitive theory that overlaps considerably with the current proposals. Further exploration is warranted given the fact that other speculators are moving in the same direction.

ACKNOWLEDGMENTS

My thanks to Lorna Champion, many of whose ideas have been included here, and to Chris Brewin and the editors for helpful comments on an earlier version of this chapter.

REFERENCES

- Abramson, L. Y., Garber, J., & Seligman, M. E. P. (1980). Learned helplessness in humans: An attributional analysis. In J. Garber, & M. E. P. Seligman (Eds.), *Human helplessness: Theory and applications*. New York: Academic Press.
- Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87*, 49-74.
- Akiskal, H. S., Bitar, A. H., Puzantian, V. R., Rosenthal, T. L., & Walker, P. W. (1978). The nosological status of neurotic depression: A prospective three-to-four-year follow-up examination in light of the primary-secondary and unipolar-bipolar dichotomies. *Archives of General Psychiatry, 35*, 756-766.
- Alloy, L. B., & Abramson, L. Y. (1979). Judgements of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology (General), 108*, 441-485.

- Alloy, L. B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review*, *91*, 112–149.
- Arieti, S., & Bemporad, J. (1978). *Severe and mild depression: The psychotherapeutic approach*. London: Tavistock.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 3, pp. 47–89). New York: Academic Press.
- Bartlett, F. C. (1932). *Remembering: A study of experimental and social psychology*. Cambridge: Cambridge University Press.
- Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: Meridian.
- Beck, A. T. (1983). Cognitive therapy of depression: New perspectives. In P. J. Clayton, & J. E. Barrett (Eds.), *Treatment of depression: Old controversies and new approaches* (pp. 265–290). New York: Raven Press.
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression: A treatment manual*. New York: Guilford Press.
- Billings, A. G., & Moos, R. H. (1984). Chronic and nonchronic unipolar depression: The differential role of environmental stressors. *Journal of Nervous and Mental Disease*, *172*, 65–75.
- Blatt, S. J., D'Afflitti, J. P., & Quinlan, D. M. (1976). Experiences of depression in normal young adults. *Journal of Abnormal Psychology*, *85*, 383–389.
- Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*, 129–148.
- Bower, G. H., & Cohen, P. R. (1982). Emotional influences in memory and thinking: Data and theory. In M. S. Clark, & S. T. Fiske (Eds.), *Affect and cognition: The seventeenth annual Carnegie symposium on cognition* (pp. 291–331). Hillsdale, NJ: Erlbaum.
- Bowers, K. S., & Meichenbaum, D. (Eds.). (1984). *The unconscious reconsidered*. New York: Wiley.
- Bowlby, J. (1980). *Attachment and Loss, Vol. 3. Loss: Sadness and Depression*. Harmondsworth: Penguin.
- Brewin, C. R. (1985). Depression and Causal attributions: What is their relation? *Psychological Bulletin*, *98*, 297–309.
- Brown, G. W., & Harris, T. (1978). *Social origins of depression*. London: Tavistock.
- Bruch, H. (1978). *The golden cage: The enigma of anorexia nervosa*. Shepton Mallet: Open Books.
- Caine, T. M., Wijesinghe, D. B. A., & Winter, D. A. (1981). *Personal styles in neurosis: Implications for small group psychotherapy and behaviour therapy*. London: Routledge & Kegan Paul.
- Champion, L. A. (1985). *A psychosocial approach to the mental health of women: Depression, social support and self-help on a new housing estate*. Unpublished doctoral thesis, University of Birmingham.
- Champion, L. A., & Power, M. J. (1987). *Cognitive approaches to depression: Towards a new synthesis*. Manuscript submitted for publication.
- Cochrane, R., & Sobol, M. (1980). Life stresses and psychological consequences. In M. P. Feldman, & J. Orford (Eds.), *The social psychology of psychological problems* (pp. 151–182). Chichester: Wiley.
- Coyne, J. C., & Gotlib, I. H. (1983). The role of cognition in depression: A critical appraisal. *Psychological Bulletin*, *94*, 472–505.
- Dixon, N. F. (1981). *Preconscious processing*. Chichester: Wiley.
- Freden, L. (1982). *Psychological aspects of depression: No way out?* Chichester: Wiley.
- Gilligan, S. G., & Bower, G. H. (1984). Cognitive consequences of emotional arousal. In C. E. Izard, J. Kagan, & R. B. Zajonc (Eds.), *Emotions, cognition and behaviour*. Cambridge: Cambridge University Press.
- Golin, S., Terrell, F., & Johnson, B. (1977). Depression and the illusion of control. *Journal of Abnormal Psychology*, *86*, 440–442.
- Hamilton, E. W., & Abramson, L. Y. (1983). Cognitive patterns and major depressive disorder: A longitudinal study in a hospital setting. *Journal of Abnormal Psychology*, *92*, 173–184.
- Hammen, C., Marks, T., Mayol, A., & DeMayo, R. (1985). Depressive self-schemas, life stress, and vulnerability to depression. *Journal of Abnormal Psychology*, *94*, 308–319.
- Henderson, A. S., Byrne, D. G., & Duncan-Jones, P. (1981). *Neurosis and the social environment*. Sydney: Academic Press.
- Hitch, G. J. (1980). Developing the concept of working memory. In G. Claxton (Ed.), *Cognitive psychology: New directions*. London: Routledge & Kegan Paul.
- Ingram, R. E. (1984). Toward an information-processing analysis of depression. *Cognitive Therapy and Research*, *8*, 443–478.

- Ingram, R. E., & Smith, T. W. (1984). Depression and internal versus external focus of attention. *Cognitive Therapy and Research*, 8, 139–152.
- Johnson, J. H., & Sarason, I. G. (1979). Moderator variables in life stress research. In I. G. Sarason, & C. D. Spielberger (Eds.), *Stress and anxiety* (Vol. 6, pp. 151–167). Washington: Hemisphere.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge: Cambridge University Press.
- Jones, W. L. (1983). *Ministering to minds diseases: A history of psychiatric treatment*. London: Heineman.
- Krantz, S. E. (1985). When depressive cognitions reflect negative realities. *Cognitive Therapy and Research*, 9, 595–610.
- Layne, C. (1983). Painful truths about depressives' cognitions. *Journal of Clinical Psychology*, 39, 848–853.
- Lewinsohn, P. M., Hoberman, H., Teri, L., & Hautzinger, M. (1985). An integrative theory of depression. In S. Reiss & R. R. Bootzin (Eds.), *Theoretical issues in behavior therapy*. New York: Academic Press.
- Lewinsohn, P. M., Larson, D. W., & Munoz, R. F. (1982). The measurement of expectancies and other cognitions in depressed individuals. *Cognitive Therapy and Research*, 6, 437–446.
- Lewinsohn, P. M., Mischel, W., Chaplin, W., & Barton, R. (1980). Social competence and depression: The role of illusory self-perceptions. *Journal of Abnormal Psychology*, 89, 203–212.
- Lloyd, C. (1980). Life events and depressive disorder reviewed: II Events as precipitating factors. *Archives of General Psychiatry*, 37, 541–548.
- Lyons, W. (1980). *Emotion*. Cambridge: Cambridge University Press.
- Mandler, G. (1984). *Mind and body: Psychology of emotion and stress*. New York: W. W. Norton.
- Mechanic, D. (1978). *Medical sociology (2nd ed.)*. New York: Free Press.
- Mechanic, D. (1986). The concept of illness behaviour: Culture, situation and personal predisposition. *Psychological Medicine*, 16, 1–7.
- Miller, G. A., Galanter, E., & Pribram, K. (1960). *Plans and the structure of behaviour*. New York: Holt, Reinhart & Winston.
- Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: The roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology (General)*, 106, 226–254.
- Nelson, R. E., & Craighead, W. E. (1977). Selective recall of positive and negative feedback, self-control behaviours, and depression. *Journal of Abnormal Psychology*, 86, 379–388.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259.
- Oakhill, J. V., & Johnson-Laird, P. N. (1985). The effects of belief on the spontaneous production of syllogistic conclusions. *Quarterly Journal of Experimental Psychology*, 37A, 553–569.
- Oatley, K., & Bolton, W. (1985). A social-cognitive theory of depression in reaction to life events. *Psychological Review*, 92, 373–388.
- Oatley, K., & Johnson-Laird, P. N. (1985, August). *Sketch for a cognitive theory of the emotions*. Paper presented at summer institute on cognitive-emotion interrelations, Colorado.
- Paykel, E. S. (1979). Recent life events in the development of the depressive disorders. In R. A. Depue (Ed.), *The psychobiology of the depressive disorders: Implications for the effects of stress* (pp. 245–262). New York: Academic Press.
- Peterson, C. (1982). Learned helplessness and attributional interventions in depression. In C. Antaki, & C. Brewin (Eds.), *Attributions and psychological change: Applications of attributional theories to clinical and educational practice* (pp. 97–115). London: Academic Press.
- Peterson, C., & Seligman, M. E. P. (1984). Causal explanations as a risk factor for depression: Theory and evidence. *Psychological Review*, 91, 347–374.
- Peterson, C., Semmel, A., von Baeyer, C., Abramson, L. Y., Metalsky, G. I., & Seligman, M. E. P. (1982). The attributional style questionnaire. *Cognitive Therapy and Research*, 6, 287–300.
- Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. Solso (Ed.), *Information processing and cognition: The Loyola symposium*. Hillsdale, NJ: Erlbaum.
- Power, M. J. (1985). Sentence production and working memory. *Quarterly Journal of Experimental Psychology*, 37A, 367–385.
- Power, M. J. (1986). A technique for measuring processing load during speech production. *Journal of Psycholinguistic Research*, 15, 371–382.

- Power, M. J. (1987). The perception of life-events in depressed inpatients and hospitalized controls. *Journal of Clinical Psychology, 43*, 206–211.
- Power, M. J., & Champion, L. A. (1986). Cognitive approaches to depression: A theoretical critique. *British Journal of Clinical Psychology, 25*, 201–212.
- Roth, D. L., & Ingram, R. E. (1985). Factors in the self-deception questionnaire: Associations with depression. *Journal of Personality and Social Psychology, 48*, 243–251.
- Rumelhart, D. E., & Norman, D. A. (1985). Representation of knowledge. In A. M. Aitkenhead, & J. M. Slack (Eds.), *Issues in cognitive modeling* (pp. 15–62). London: Erlbaum.
- Sacco, W. P., & Beck, A. T. (1985). Cognitive therapy of depression. In E. E. Beckham, & W. R. Leber (Eds.), *Handbook of depression: Treatment, assessment and research*. Homewood, IL: The Dorsey Press.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: 1. Detection, search and attention. *Psychological Review, 84*, 1–66.
- Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgements of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45*, 513–523.
- Seligman, M. E. P. (1975). *Helplessness: On depression, development and death*. San Francisco: Freeman.
- Seligman, M. E. P. (1981). A learned helplessness point of view. In L. P. Rehm (Ed.), *Behaviour therapy for depression: Present status and future directions*. New York: Academic Press.
- Seligman, M. E. P., Abramson, L. Y., Semmel, A. M., & von Baeyer, C. (1979). Depressive attributional style. *Journal of Abnormal Psychology, 88*, 242–243.
- Shiffrin, R. M., & Schneider, W. (1984). Automatic and controlled processing revisited. *Psychological Review, 91*, 269–276.
- Smith, T. W., & Greenberg, J. (1981). Depression and self-focused attention. *Motivation and Emotion, 5*, 323–331.
- Teasdale, J. D. (1983). Negative thinking in depression: Cause, effect, or reciprocal relationship? *Advances in Behaviour Research and Therapy, 5*, 3–25.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*, 548–573.
- Wilkinson, I. M., & Blackburn, I. M. (1981). Cognitive style in depressed and recovered depressed patients. *British Journal of Clinical Psychology, 20*, 283–292.

CHAPTER 12

Cognitive Treatment of Depression

J. M. G. Williams

INTRODUCTION

Something over 150 different psychotherapies for emotional disturbances exist. In the light of this it is important to retain perspective in evaluating the significance of any new therapy. Cognitive-behavior therapy deserves our attention because its efficacy has been studied more systematically than most forms of psychotherapy for clinical depression. Not only has it been compared with other forms of psychological treatment (e.g., psychodynamic approaches and relaxation; McLean & Hakstian, 1979) but also compared with the most commonly used physical treatment: antidepressant medication, (Blackburn, Bishop, Glen, Whalley, & Christie, 1981; McLean & Hakstian, 1979; Murphy, Simons, Wetzel, & Lustman, 1984; Rush *et al.*, 1977).

Cognitive therapy for depression as it is most widely practiced was pioneered in the United States by the psychiatrist A. T. Beck. Trained in psychoanalytic methods, Beck became more interested in the depressive content of patient's cognitions—thoughts, images, memories; in the self-defeating nature of many patients' assumptions; and in the way in which patients tended to “screen out” positive information, encoding only negative interpretations of otherwise neutral situations. Beck's theory is most clearly laid out in his 1976 book *Cognitive Therapy and the Emotional Disorders*. In that book three main components of a theory of emotional disorders are outlined. The first component is the presence of negative automatic thoughts—automatic by virtue of their coming out of the blue, often seemingly unprompted by events and not necessarily the results of directed thinking. They seem immediate and often valid in the sense that they are often accepted unchallenged by the recipient. Their effect is to disrupt mood, and to cause further thoughts to emerge in a downward thought-affect spiral. Depressive thoughts can be characterized in terms of a *cognitive-triad*—a negative view of the self (e.g., “I'm a failure”), the world (e.g., “this neighborhood is a terrible place”) and the future (e.g., “everything will turn out badly”).

The second component is the presence of systematic logical errors in the thinking of depressed individuals. Several categories (not mutually exclusive) have been distinguished: *arbitrary inference* (e.g., someone concludes that a friend has fallen out with them because they did not smile at him or her); *overgeneralization* (e.g., “failure on this exam means I’ll never pass the other exams”); *selective abstraction* (e.g., when people only notice the few bad things in reports about themselves); *magnification and minimization* (e.g., when a person exaggerates the effect of a negative event (catastrophizes) or minimizes the impact of a positive event); *personalization* (when people attribute bad things to themselves despite evidence to the contrary) *dichotomous thinking* (all or nothing, black/white thinking, e.g., “only a miracle can make me well again,” or “if he leaves me, I may as well be dead”).

The third component of the cognitive model is the presence of *depressogenic schemata*. These general, long-lasting attitudes or assumptions about the world represent the way in which the individual organizes his or her past experience, and is suggested to be the system by which incoming information about the world is classified. This is one of the earliest concepts in Beck’s theoretical writings. In 1964 he defined the schema as

a structure for screening, coding and evaluating impinging stimuli. In terms of the individual’s adaptation to external reality, it is regarded as the mode by which the environment is broken down and organized into its many psychologically relevant facets; on the basis of the matrix of schemas, the individual is able to orient himself in relation to time and space and to categorise and interpret his experiences in a meaningful way. (p. 564)

According to the theory, depressive schemata develop over many years and, although they may not be evident in later life, remain ready to be activated by certain combinations of stressful circumstances.

Although Beck himself had been writing about his cognitive theory of depression since the 1960s (e.g., Beck, 1964, 1967), it was not until the middle and late 1970s that worldwide interest was shown in the therapeutic techniques he and his clinic were advocating. There are probably many reasons for this increased interest, but it is worth mentioning just two that probably had more influence than most.

Firstly, psychological treatments for anxiety neuroses and phobic states, which had been revolutionized by the introduction of the behavioral techniques of desensitization and flooding, had started to move toward more cognitive approaches. Many psychologists had followed the lead of such clinical researchers as Meichenbaum (1977), who had confirmed the efficacy of therapies for anxiety that attempted to change patients’ “self-talk.” Through such work, clinicians were introduced to the notion that a patient’s “private commentary on the world” was an important determinant of his or her reactions to ambiguous (but potentially threatening) situations.

Secondly, a cognitive theory of depression had been proposed by Seligman (1975) in his concept of learned helplessness. Developed out of the observations of animals’ failure to escape and avoid traumatic shock, Seligman and his colleagues (Maier & Seligman, 1976) suggested that animals and humans could, under certain circumstances, learn that responding and outcome are independent. This cognitive evaluation: “nothing in the response repertoire makes any difference to the outcome” was shown to produce motivational, cognitive, and emotional deficits, and suggested

as an adequate (sufficient though not necessary) model of reactive depression. Although the model has since gone through many changes, its main impact in the mid-1970s was to give psychologists a plausible learning-based model for some depressive phenomena. But it was a model with few therapy implications. Seligman (1975) and Abramson, Seligman, and Teasdale (1978) had made some suggestions as to what sort of treatments ought to work (e.g., success experiences, reattribution) but these implications were not worked out in detail. Psychologists had a learning-based cognitive model, but no learning-based cognitive treatment. It was natural to search for a therapy that could address some of these issues directly—and the cognitive techniques of Beck were there to fill the void.

Cognitive therapy involves a balance between behavioral techniques (such as task assignment) and cognitive techniques (such as cognitive rehearsal). As the patient progresses through therapy, the balance shifts away from behavioral to cognitive techniques. The more severe the patient at the outset, the greater the emphasis on behavioral techniques. If a patient is only moderately depressed at the outset, then the therapy starts with a mixture of cognitive and behavioral techniques. Within both the behavioral and cognitive components the progression is from the simple to the more complex. Thus behavioral techniques shift from graded task assignment, where therapist and patient decide how to split up tasks so that only parts of tasks are set as goals for between-session homework, to task assignment, where whole tasks are assigned. Similarly, cognitive techniques shift in emphasis from addressing cognitive events (simple thoughts, images, memories) to addressing underlying assumptions and cognitive schemata. (For further details of specific techniques, see Beck, Shaw, Rush, & Emery, 1979; Williams, 1984a). Cognitive therapy is normally limited to 15 to 20 sessions of 50 minutes each, once weekly, though for more severely depressed patients sessions are held twice weekly for the first 4 to 5 weeks. It is interesting to note that in the study by Blackburn *et al.* (1981), patients were considered nonresponders if their Beck Depression Inventory (Beck, 1967) or Hamilton Scores (Hamilton, 1967) had not dropped by 50% after 12 weeks of treatment. Sometimes this improvement can be quite rapid (Teasdale, Fennell, Hibbert, & Amies, 1984).

Over the last decade there has been a rapid expansion of the use of the techniques that Beck had drawn together. Two questions have dominated the research literature: Do they work? and, if so, What are the effective components?

The first question will not be dwelt on here. Elsewhere, I have reviewed some of the recent outcome literature and examined the extent to which it meets various criticisms (Williams, 1984b). A very crude overview of outcome is shown in Table 1, which gives an updated summary of studies on depressed patients (i.e., excluding volunteer depressed subjects) that have used the Beck Depression Inventory (Beck, 1967) as an outcome measure. Most find that cognitive therapy with or without drugs is as effective as drugs alone, though Blackburn's results suggest that, especially for outpatient depressives, the combination of drugs and cognitive therapy can be maximal. At the very least it is possible to conclude from these data that drugs and cognitive therapy do not inhibit each other. The study by Murphy *et al.* (1984) is interesting in that they found no difference between drug and cognitive therapy. As

TABLE 1. Studies Using Beck Depression Inventory—Changes with Treatment

Study	Treatment	Subjects	BDI score		Proportionate change
			Pre	Post	
Rush <i>et al.</i> (1977)	CBT ^a	Outpatients	30.3	5.9	0.81
	Drug ^c	Outpatients	30.8	13.0	0.58
McLean & Hakstian (1979)	CBT	Outpatients	26.8	9.7	0.64
	Drug ^d	Outpatients	27.2	14.1	0.48
	Relaxation	Outpatients	26.8	15.0	0.44
	Psychotherapy	Outpatients	27.0	16.8	0.38
Blackburn <i>et al.</i> (1981)	CBT	GP patients	—	—	0.84
	CBT + drug ^b	Outpatients	—	—	0.79
	CBT + drug ^b	GP patients	—	—	0.72
	Drugs ^e alone	Outpatients	—	—	0.60
	CBT	Outpatients	—	—	0.48
	Drugs ^b alone	GP patients	—	—	0.14
Murphy <i>et al.</i> (1984)	CBT	Outpatients	28.7	9.5	0.67
	Drug ^c	Outpatients	29.2	8.9	0.69
	CBT + Drug	Outpatients	29.1	8.8	0.70
	CBT + Placebo	Outpatients	30.3	8.2	0.73
Teasdale <i>et al.</i> (1984)	CBT + Routine treatment	GP patients	30.0	8.0	0.73
		GP patients	29.0	18.5	0.36
	Routine treatment				
Beck <i>et al.</i> (1985)	CBT	Outpatients	31.0	8.6	0.72
	CBT + Drug ^d	Outpatients	30.0	10.0	0.67

^aCBT = Cognitive behavior therapy

^bAmitriptyline and clomipramine.

^cImipramine

^dAmitriptyline

^eNortriptyline

can be seen from Table 1, this is partly because the drug group responded better in this study than in others and partly because the cognitive therapy group did less well. This may be the result of choosing nortriptyline as the drug for the study and taking regular blood tests to ensure that the drug levels in the patients' bloodstreams were within the therapeutic window. The care with which the drug group were treated may explain their better response in this study compared with other studies. It is to the second major question of which processes affect outcome that we now turn.

There are a number of strategies available to investigators who wish to elucidate the factors underlying response to therapy. Firstly, there are experiments that take different subcomponents of treatment and examine their effectiveness on a range of variables. Secondly, one may examine individual differences in pretreatment characteristics, intrapersonal and interpersonal variables that differentially predict response to an entire treatment package. Thirdly, one may measure the progress of some process variables throughout treatment. These three approaches are not, of course, mutually exclusive. Each can be done within the same outcome trial. Nevertheless they are conceptually separable, and we shall consider them each in turn, although in some cases the same study may be relevant to more than one issue.

ANALYSIS OF SUBCOMPONENTS OF TREATMENT PACKAGE

Within this category there are two types of research that have been done. The first, using volunteer depressed subjects has contrasted a small number of behavioral techniques with a small number of cognitive techniques (Taylor & Marshall, 1977; Wilson, Goldin, & Charbonneau-Powis, 1983; Zeiss, Lewinsohn, & Munoz, 1979). The second, using depressed patients, has attempted to examine the impact of one or two specific therapeutic techniques in isolation (Fennell & Teasdale, 1984; Rotzer, Nabitz, Koch, & Pflug, 1982).

Volunteer Studies - Behavioral versus Cognitive Packages. Taylor and Marshall (1977) recruited through advertisement 45 subjects with BDI scores greater than 13 and whose self-reported depressed mood has lasted at least 2 weeks (mean BDI = 21.2, "moderately severely depressed"). Subjects were randomly assigned to one of four groups: cognitive only, behavioral only, cognitive and behavioral combined, and waiting-list control. The rationale for the cognitive therapy given was that "depressed mood is rooted in self-evaluation." Subjects were taught how to become aware of thoughts that occurred between an event and consequent affective disturbance, and instructed to use alternative self-statements to cope with such situations when they occurred. With a therapist's help, a list of positive self-statements was constructed, and subjects were taught to read through the list before engaging in a high probability behavior (*coverant control*, using the Premack principle of making a high probability behavior contingent on a low probability behavior).

The rationale for the behavioral treatment was that "depression results from insufficient positive reinforcement." Subjects were given help in identifying situations that produced depressed mood, and in learning new alternative patterns of behavior. Role play, modeling and homework assignments to rehearse new techniques were used, often with the aim of promoting more assertive, socially skilled behavior. The combined treatment would have made it impossible to spend more than half the time (on average) on each component. Despite this, the results of the six 40-minute sessions over 4 weeks showed a clear superiority for the combined treatment over each one alone, which in turn were superior to no treatment (assessed by BDI, MMPI-D and visual analogue mood scales). The trend of these results was still clearly visible on follow-up assessment 5 weeks later.

This study is important because it was the first to address directly the issue of which elements in a cognitive-behavioral package are responsible for therapeutic progress. The finding that integrating behavioral into a cognitive context (and *vice versa*) works better than spending the full time engaging in one or the other model, has direct therapeutic implications if work with depressed patients bears it out.

Zeiss *et al.* (1979) recruited subjects on the basis of an MMPI(D) score greater than 80 and a structured interview. Subjects were allocated to one of four groups: cognitive therapy, interpersonal skills training, pleasant activity scheduling, or waiting list. In addition to using MMPI(D) scores as a general outcome measure, 7 hours of comprehensive assessment procedures were implemented every month (four occasions in all) to assess the specific subcomponents purported to mediate the efficacy of the various therapies. Thus three ratings (including one observer rating) of interpersonal behavior were taken; four ratings (including an observer rating) of cognitive

style were taken; and the Pleasant Events Schedule was used to assess frequency and subjective pleasantness of activities performed. Results showed no differences between the three active therapy groups, though they all produced more improvement than the waiting list-controls.

Wilson *et al.* (1983) recruited 25 subjects between 20 and 60 years of age who had a score of at least 17 on the Beck Depression Inventory and self-reported experiences of frequent episodes of depression in the previous 3 months. These were randomly allocated to behavior therapy (8 weeks), cognitive therapy (8 weeks) and waiting-list control. The behavior therapy was based on Lewinsohn's model (Lewinsohn, 1976), the overall aim being to increase the frequency, quality, and range of activities through the use of activity schedules. Cognitive therapy excluded behavioral assignments, concentrating instead on negative cognitive distortions (logical errors, such as overgeneralization and selective abstraction) and irrational beliefs (e.g., "I must be perfect at everything I do"). Frequent negative thoughts and infrequent positive thoughts were identified within sessions and instructions given to employ the positive counterstatements whenever the subjects became aware of negative thoughts. Results showed that both treatments produced significant and substantial improvement over 8 weeks of therapy compared with the waiting-list control. Few differences were found between the cognitive and behavioral treatments. Both types of treatment produced the same degree of increase in pleasant activities at midtreatment and at posttreatment, but the cognitive therapy produced a significantly greater change in the self-rated frequency of positive cognitions at midtreatment than did the behavior therapy. This gap had closed, however, by posttreatment and follow-up (5 months) assessment. The authors ignore a potentially important aspect of their data which is consistent with the above finding for positive cognitions. The overall percentage change score from pre- to posttreatment in Beck Depression Inventory was 65% for the behavioral and 67% for the cognitive. However, if one examines the BDI scores after just 4 weeks of therapy, the percentage change for behavior therapy is 22% and for cognitive therapy is more than double at 49%. It appears that although both therapies have the same end-point, they may have had different recovery rates that a finer-grained analysis of the time course of different process variables might have shown. The cognitive therapy package appears to work fastest in the initial stage of therapy, though considerable caution must be exercised in generalizing from these volunteer depressed subjects to clinically depressed individuals who may be less accessible to cognitive procedures in the early phase of therapy.

Note that neither Zeiss *et al.* (1979) nor Wilson *et al.* (1983) used a combined behavioral and cognitive experimental group, so were unable to test the replicability of Taylor and Marshall's conclusion that combination treatment exceeds either treatment alone in its effects.

Of course, there is a "labeling" problem with research that compares cognitive with behavioral techniques: When does a technique change from being behavioral and become cognitive? Could not any increase in efficacy of a combined treatment, wherever found, be because the therapist had a wider range of techniques from which to choose, so could match individual treatment components with the individual characteristics of the patient?

A second problem is that of interpreting what the nature of the behavioral versus cognitive difference is. I know of no behavioral therapist who has not been

able to translate so called cognitive techniques into behavioral terms, and cognitive therapists are well-accustomed to arguing that the crucial aspect of behavioral techniques is the cognitive changes they bring about. In part the difficulty arises because of the philosophical issue that behavioral and cognitive languages may be interchangeable because they belong to different universes of discourse. But that is not to say there will not be occasions on which the context will demand one sort of language rather than another. For example, devising an intervention program to deal with a child who appears depressed and withdrawn as the result of poor institutionalization or inadequate parenting, or an intervention program for chronically depressed and retarded adult depressives, may demand a careful functional analysis concentrating on behaviors. By contrast if one's assessment finds a patient complaining about their poor memory and concentration, it will be helpful to have a fully elaborated cognitive model for how such problems may arise and how they may be remediated (e.g., Watts & Sharrock, 1985). Furthermore, contact with those researchers in the field of cognitive modeling, whose findings may have much to contribute to the understanding of depression (see chapter by Power) is likely to be much impoverished unless one has available a cognitive language in which to describe clinical depressive phenomena. In summary, differentiating between behavioral and cognitive concepts is likely to be useful only where the clinical or research context demands explicit use of one or other model. Looking back at the outcome literature that has compared behavioral and cognitive techniques one may see that this differentiation is inadequate because it attempts to apply different techniques in the same clinical and research context, that is, where distinguishing between behavioral and cognitive concepts is least likely to be useful.

We must not let these arguments, however, obscure a very important point about the outcome research comparing behavioral and cognitive techniques. Whatever the label, it has been (and will continue to be) useful to contrast some techniques with some other techniques in the treatment of carefully matched comparison groups. There will always remain the question of what actual techniques, however formulated, are most effective for which patient. It is to studies that attempt to address this question that we now turn.

Patient Studies: Specific Components of Therapy. Rotzer *et al.* (1982) examined the therapeutic effectiveness of Activity Training alone versus Activity Training plus Self-Regulation Training versus Activity Training plus Coverant Control. Thirty-eight depressed outpatients were randomly assigned to receive either one of three treatments. Each treatment lasted for 12 weeks. Half the patients received medication (amitriptyline or maprotiline up to 150 mg/day) in addition to the psychological treatments, whereas half did not. The characteristics of the subjects in this study show them to be a more appropriate group than that in the volunteer studies reported earlier. The average time since first onset of depression was 8 years 6 months, and the current illness phase had lasted for a mean of 32 months. Sixteen of the sample had been chronically depressed for 10 years or more; nine of the sample were endogenous (ICD 296.1) and 29 neurotic/reactive (ICD 300.4).

Results on all measures in the study showed significant improvement in all treatments, with or without drugs. The blind Hamilton Rating Scores confirmed these results, the Activity Training group changing from a mean of 19 (pre) to 6 (post); Activity Training plus Self-Regulation from 16 (pre) to 10 (post); and Activity

Training plus Coverant Control from 15 (pre) to 5 (post). Thus, the overall percentage change scores from the three groups' HRS of 65%, 38%, and 67% respectively compare fairly well with the percentage change scores on the HRS in the Blackburn *et al.* (1981) Edinburgh study, the slightly lower overall change in the Rotzer *et al.* patients being perhaps attributable to the reduced treatment length (12 weeks, versus up to 20 weeks in the Edinburgh Study). The conclusion of their study is that Self-Regulation Training and Coverant Control do not add any therapeutic benefit over and above Activity Training.

Fennell and Teasdale (1984) examined a different subcomponent of cognitive behavior therapy and in a different way. They studied the effect of distraction by using short periods of "mini-therapy." Sixteen depressed patients who met research diagnostic criteria for primary major depressive disorder (Spitzer, Endicott, & Robins, 1978) and who scored > 20 on the Beck Depression Inventory were given 5 minutes of distraction (concentrating on and describing aloud colored slides of outdoor scenes) or 5 minutes of a control condition (sitting quietly looking at a square of white light projected on the wall) in a within-subject design. The distraction condition significantly affected the frequency of negative thoughts, which in turn affected self-rated depressed mood, as well as two psychomotor measures, counting speed and writing speed. Interestingly, these effects were shown most clearly by the patients who fell at the low-endogenous end of the Newcastle Diagnostic Scale (Carney, Roth, & Garside, 1965). A similar result had also been found by Teasdale and Rezin (1978) and by Davis (1982; cited in Williams, 1984b).

What conclusions may be drawn from these analyses of the subcomponents of the treatment packages? Firstly, that used alone, cognitive and behavioral procedures appear to be equally effective, but that their combination may well be more effective. The Rotzer *et al.* clinical study is the one investigation which runs contrary to this conclusion. It is noticeable that their Activity Training Alone group actually had an unusually large response for such a chronic group (compare Teasdale & Fennell, 1984; Harpin, Lieberman, Marks, Stern, & Bohannon, 1982), and that many cognitive therapists report that encouraging any sort of activity increase is a considerable struggle with many chronically depressed patients. More clinical research is clearly needed to examine which combinations are most effective, and for which patient.

A further conclusion that may be drawn is that if given a treatment package, different patients will tend to respond to different aspects of the therapy. That this is so can be inferred from Fennell and Teasdale's (1984) results showing that distraction works for the less endogenous patients. We may place this result in the context of results from outcome trials (Blackburn *et al.*, 1981; Murphy *et al.*, 1984) that show that endogeneity of depression does not predict overall response to the entire treatment package. But if different patients do respond to different aspects of the package, how can one determine the processes that are underlying effectiveness? There are simply too many combinations (19 separate techniques can be combined in over 2 million different ways!) to allow straightforward conclusions. Analyzing one or two components one by one is possible, though a very long-term enterprise. But more importantly, it may not reveal what the important combinations are. Let us see how far the other two strategies (studying individual differences in treatment response, and examining process variables) may help in this respect.

PREDICTORS OF TREATMENT RESPONSE

Many investigators have looked for individual difference variables that would predict treatment response, but without a great deal of success. Neither age nor I.Q. predicts outcome, nor does endogeneity. McLean and Hakstian (1979) measured a number of possible predictor variables (including personality-EPQ), in a study of 178 moderately clinically depressed clients. Their conclusion was that their results showed "somewhat surprisingly, that none of the variables often thought to predict a treatment outcome—age, severity, number of previous episodes, personality measures, and so forth—are very powerful predictors" (p. 835) and refer to other authors who have concluded that "the relation between personality and depression in predicting outcome is elusive" (Becker, 1974; Chodoff, 1972). More recently, however, there has been some small progress made in finding predictor variables. Murphy *et al.* (1984) found that high scores on Rosenbaum's Self-Control Scale (Rosenbaum, 1980) predicted response to cognitive behavior therapy. This scale purports to assess "learned resourcefulness," the tendency for an individual to use a style of coping that emphasizes self-reliance, which correlates with internality as measured by Rotter's Locus of Control Scale. Interestingly, patients in the antidepressant drug group (nortriptyline) showed a better response to the medication if they had low scores on this scale. There was then an interaction between natural coping style and type of treatment—whether the treatment was consistent or inconsistent with the patient's preference for different sorts of coping strategy.

An outcome trial using G.P. patients who met the criteria for major depressive disorder (Spitzer *et al.*, 1978) by Teasdale *et al.* (1984) has also identified some predictors of outcome. As part of the cognitive therapy trial, all patients were given a booklet "Coping with Depression" at the end of the first session, and asked to read it as their homework for the next session. The booklet gives details of the treatment as well as describing some symptoms of depression in cognitive terms. During their second session, patients' comments on how helpful they had found the booklet were recorded. These patients' reactions were independently rated for how helpful and applicable the booklet had seemed to be. These ratings, taken on Week 2, predicted outcome after 12 weeks significantly (Fennell & Teasdale, 1987). What could explain the predictive power of this early reaction? Firstly, perhaps, response to the booklet reflected general motivation to participate in the therapy. Secondly, a good response could reflect understanding and sympathy with the cognitive model (irrespective of motivation). In this case the conclusion would be that cognitive therapy is most appropriate if it meshes with an individual's own understanding of his symptoms. This would of course be consistent with the Murphy *et al.* findings regarding the self-control scale. Perhaps those who see some hope in the cognitive behavioral model are those who naturally prefer coping by action rather than inaction. Finally, the response to the booklet may reflect an important individual difference in patients' abilities to distance themselves from their own symptoms. Most behavior therapists or cognitive therapists are aware that a sign of progress in therapy is when a patient reports that they are able to see their problems with distance or perspective. They may still become upset by an event, for example, but they can see that it is a natural reaction that will wear off and are prepared to be more passive in observing the

time-course of the recovery process. It is interesting to speculate that subjects who, in reading a booklet early in therapy, are able to see how they fit into this more general picture, may be demonstrating an early ability to distance themselves from their depressive thoughts, affect, and behavior.

PROCESS RESEARCH

Attempts to assess not only outcome but also process variables may be exemplified by reference to Zeiss *et al.* (1979), who examined the effect of different therapies on measures that were designed to be specifically sensitive to one of the treatment components, and Simons, Garfield, and Murphy, (1984), who examined the effect of drugs and cognitive therapy alone and combined on the Dysfunctional Attitude Scale (which should be most affected by cognitive procedures).

The details of the Zeiss *et al.* trial have been given earlier. The different ratings of interpersonal behavior, cognitive style, and pleasant events were designed as specific checks on progress of the interpersonal skills, cognitive, and pleasant activity scheduling therapies respectively. However, all assessments improved equally in the three different types of treatments. It was not the case that each therapy had its effect by modifying the particular mediating variable purported to underlie its particular set of therapeutic strategies. Rather, the authors suggest "the label of therapy does not ensure that the behaviors labelled will be those most directly affected," and conclude that psychological treatment may be producing "nonspecific" improvement no matter what treatment model is being tested.

There is little doubt that nonspecific treatment effects are present in studies of the effects of cognitive-behavior therapy but this study may not constitute as strong evidence for their involvement as they suppose. Firstly, the results may not be generalized to clinically depressed individuals whose motivation is not perhaps as consistently positive as these recruited volunteers. Secondly, the assessments of the process variables were only made once a month, hardly frequent enough to justify such a general conclusion. The time scale over which cognitive changes, if present, would be expected to affect mood and behavior would be hours and days, rather than days and weeks, and this could be said to be true of the effect of interpersonal skills training and activity scheduling on the other process variables. This makes the once-a-month monitoring seem very inadequate. This is because it assumes that wholesale changes in one variable will occur prior to wholesale changes in the next. It is more likely that change will be interactive, with small changes in one variable predicting small changes in the next, and so on.

Unlike Zeiss *et al.* (1979), who were concerned with whether cognitive techniques had a similar mode of action as behavioral techniques, Simons *et al.* (1984) were concerned with comparing the effect of antidepressant medication on the mechanisms that are supposed to mediate cognitive behavior techniques. Simons *et al.* examined changes in scores on the Dysfunctional Attitude Scale during the course of the therapy trial reported by Murphy *et al.* (1984). This scale purports to measure maladaptive assumptions by asking patients how much they agree with such statements as "My value as a person depends greatly on what others think of me" and "If people whom

I care about reject me it means there is something wrong with me.” They found that reductions in mean DAS scores for the drug group exactly paralleled those from the cognitive therapy group. It seemed that if a treatment is effective, it changes a person’s cognitions, whether these are the focus of treatment or not. If this is the case, two conclusions would seem to follow. Firstly, that one may as well give drugs if they are going to have identical effects on the mediating mechanisms. Secondly, one may doubt the importance of cognitions as a causal factor. Let us consider each of these in turn.

One must be careful in jumping to the conclusion that drugs are equally appropriate for all patients. Firstly, for some people (e.g., those with heart problems) antidepressant medication has a physical risk factor to be taken into account, especially if there is a risk of overdose. Secondly, some patients will have had many antidepressants without effect, and one needs a psychotherapy of proven effectiveness to be available as an alternative. Thirdly, just as some patients insist their problems are biochemical and demand medication, so others insist their problems are psychological and refuse drugs. Management procedures are needed for these patients too. But finally, there is now evidence that although drugs produce an average score for dysfunctional attitudes that is within normal nondepressed limits, if one examines individual items there are some peculiarly persistent dysfunctional attitudes that seem to remain after the patient is otherwise symptom free (Reda, 1983). Reda studied 60 patients (DSM-III “Major depressive disorder” with melancholia) who had been treated with amitriptyline at 75 mg to 120 mg per day then discharged on a maintenance dose of 50 mg per day. They were rated blind on the Hamilton Rating Scale (Hamilton, 1967) and on a scale that assessed dysfunctional attitudes at discharge and (for half the sample) one year after discharge. Although there was no overall significant difference between dysfunctional cognitions at discharge and follow-up in the recovered depressed group and matched normal control (consistent with the findings of Simon *et al.*, 1984) some residual maladaptive beliefs continued to differentiate the groups (see Table 2).

So it may not be so straightforward as simply looking at mean change scores on these cognitive variables. One must discriminate the dimensions on which change occurs.

The second doubt arising from recent results is whether cognitions are causal in depression. Many people have argued that cognitions do play a causal role (Beck, 1976), others that they do not (Phillips, 1981). The present author believes that the debate about causality is philosophically interesting but clinically useless. If cognitions were one day established as the cause of some depressions, it would not mean that behavioral or physical aspects were irrelevant for those depressions because these other factors may well need to be manipulated to break the vicious circle. Similarly, if biological factors were one day established beyond doubt as the cause of some other subtypes of depression, one may still require a psychological, functional analysis and intervention as the point of entry. What we need are adequate *functional* models, not causal models—that is, we need to know what effect modifying one part of the system has on other parts of it. To do that, we need adequate tools to make specific assessments of the components of the system. It is at this point that we need to ask what cognitive psychology has to offer.

TABLE 2. Beliefs That Persisted after Recovery from Depression Treated with Amitriptyline

DAS item	Sample
I feel well only when I have complete control over the situation.	M & F
Turning to someone else for advice or help is an admission of weakness.	F
If you cannot do something well, there is little point in doing it at all.	F
People should have a reasonable likelihood of success before undertaking anything.	M
I should be able to please everybody.	M & F
It is shameful for a person to display his weakness.	M
If a person is not a success, then his life is meaningless.	F
If I do not do as well as other people, it means I am an inferior human being.	F
I should always have complete control over my feelings.	M & F
If I fail at my work, then I am a failure as a person.	F
My happiness depends more on other people's judgment than it does on me.	F
I ought to be able to solve my problems quickly and without a great deal of effort.	M & F
I always see the negative aspect of everything.	M & F

Note: M & F = shown by both male and female recovered patients; M = shown by male recovered patients; F = shown by female recovered patients.

COGNITIVE PSYCHOLOGY AND DEPRESSION

Several authors have suggested recently that experimental cognitive psychology offers clinicians methods with which they can evaluate the subcomponents of the cognitive system (Kihlstrom & Nasby, 1981; Merluzzi, Rudy, & Glass, 1981; Williams, 1984a). Cognitive psychologists have developed methods of examining the logical errors people make in thinking, the biased way in which information may be selected, stored, and retrieved, and the way in which information is schematized, which affects many cognitive subsystems. What is significant about much of this research, however, is that their application to depression has been largely post hoc with respect to the development of clinical aspects of the cognitive therapy model. It is worth pausing for a moment to reflect on this.

This chapter has up until now mainly been concerned with examining the question of what the important effective elements in cognitive behavior therapy are. That is, it has started from the findings that cognitive therapy is effective in treating at least some subtypes of depressed patient, and worked backwards to ask why this may be so. This is a very different approach from starting with theories developed in the laboratory and then applying them to clinical states, which was more characteristic of behavioral approaches to anxiety in the 1960s. (The only main cognitive theory of depression developed in the laboratory, that of Seligman, did not, as we have already observed, have many specific and readily applicable treatment applications.) The structure of the chapter reflects the way in which the research itself has developed.

Beck's theories were derived from clinical observation, and it has been left to other clinical researchers to explore whether the processes that Beck inferred to explain his observations actually do exist. Some aspects of this research have been reviewed by Teasdale (1983) and by Williams (1984a, Chap. 8) and will not be reviewed again here. But it can be seen how the theoretical foundations of cognitive therapy have

been elaborated and made more explicit in response to rather than prior to clinical observation and intervention. In the light of this, what is the role of experimental cognitive research for cognitive therapy?

One obvious role is the development of more specific methods of assessing cognitive change as a result of treatment. To date, most of the outcome studies on anxiety and depression have relied on either observational measures by therapists or independent assessors, or self-report by the subject. Behavioral avoidance tests have also been used in treatments for anxiety but analogous behavioral tests are not readily available for assessing depressed patients. It is an assumption of cognitive therapy as practiced by Beck and his colleagues that until the underlying dysfunctional attitudes and biased schemata have been treated, the patient will remain vulnerable to react to future life stresses by becoming depressed. Self-report measures of cognitive vulnerability have generally been unsuccessful in picking up evidence of such vulnerability between episodes of depression, but this may be because the measures are too gross, being mainly tests of people's explicit responses to explicitly valent stimuli. More direct measures of underlying schemata are now available in the form, for example, of recall measures following the self-referential encoding of personally relevant trait adjectives (Kuiper & McDonald, 1982) and in developments of the use of emotional stroop task (Gotlib & McCann, 1984; Mathews & MacLeod, 1985; Watts, McKenna, Sharrock, & Trezise, 1986; Williams & Broadbent, 1986a; Williams & Nulty, 1986c). This research is in its infancy at the moment, but it shows promising signs that in a few years' time methods will be available which will validly and reliably assess underlying cognitive biases.

A second purpose for cognitive research in this area is to examine what the important core features of the cognitive/appraisal schemata of depressed patients are. In examining Beck's 1976 theory early on in this chapter, we saw how many of the logical errors that he identifies are not mutually exclusive. What is unclear at the present time is whether one needs to employ so many different categories of biased thinking or whether there may be one or two core errors that need to be made explicit and investigated more closely and from which all the others follow. A prime candidate for such a core construct is that of overgeneralization (Carver & Ganellan, 1983; Teasdale, 1983).

But a major usefulness of cognitive research will remain in elucidating fundamental ways in which cognitive processes contribute to the onset, maintenance, and alleviation of emotional disturbance. Let us look more closely at two examples of research that attempt to examine these processes.

Strack, Schwarz, and Gschneidinger (1985) examine the effect of memory on judgments of current well-being. They start from the observation that objective circumstances, although important, often explain only a small part of the variance in subjective ratings of happiness and satisfaction. When people consider the quality of life, events can only therefore be one part of a much more complex story that must also take account of the psychological mechanisms mediating between the event and the subject's well-being. In a series of studies they asked students to recall positive or negative events either from their present life or from their past. They found that recall of positive events from the present had a positive effect on ratings of subjective well-being. Recall of negative current events had the opposite effect. However, events recalled from the past tended to have a contrast effect on current well-being, recalling

past negative events enhancing well-being and recalling past positive events damaging it. When does recalling a valent past event have a congruent effect on well-being, and when does it have the contrast effect they observed? The authors noticed that in their first experiment, the present recall condition tended to have a small effect on current mood, whereas no such mood effects were observed at all for past recall. In a second experiment they confirmed that if the past event was recalled vividly enough to disturb current mood, then the judgment of well-being no longer showed a contrast effect but moved in a direction congruent with the valence of the event being recalled. They replicated this finding in a third experiment in which they manipulated the manner in which the past event was recalled. If subjects were asked to recall a positive or negative past event in the context of asking why it had occurred, their recall was less detailed and vivid than if they recalled it in the context of asking how it had occurred. In the why (nonvivid) condition, past positive events caused decreases and past negative events an increase in subjective well-being. In the how condition a mood congruent effect of recalled event on current well-being was observed.

The general conclusion that the effects of recalling past events on current well-being depends on whether they affect current mood is an important one, because it suggests remedial strategies for some patients that might increase their resistance to fluctuations of current mood when memories are causing a problem. But if the therapist decides to work on the memories themselves, this research suggests that the way in which the memory is recalled will have an important bearing on its affective impact.

There are of course important caveats to be pointed out. These subjects were not depressed so their current mood was presumably not vulnerable as one might expect a depressed person's to be. Secondly, the recalled life events were unlikely to be as ego-threatening as are those of many depressed patients. Nevertheless, the research serves as a good model for how research on the interface between cognitive appraisal and more basic memory processes may proceed.

The other illustration of research on basic cognitive processes comes from some of our own work. In Cambridge we have been following up work started at the Maudsley Hospital in London by Lloyd and Lishman (1975). They found biased retrieval in the memories of depressed patients, such that the more depressed individuals had greater difficulty in remembering a pleasant event from his or her past, but remembered negative events very easily. Lloyd and Lishman speculated that this may be an example of state-dependent learning in which an organism remembers better what it has learned if returned to the place or state of mind in which the original learning took place. Teasdale and Fogarty (1979) used mood induction procedures with nondepressed subjects and found that similar retrieval biases can be produced, demonstrating that it is current mood that is responsible for the bias, rather than actually having more negative and less positive events in one's life history to sample from.

Williams and Broadbent (1986b) studied patients who had attempted suicide by overdose. In an autobiographical memory test, patients were asked to try and remember specific personal events in their life (recent or remote, important or trivial). Patients' memories were cued with pleasant or unpleasant emotion words (e.g., "happy," "safe," "sorry," "angry"). Compared with matched control patients, the emotionally disturbed sample were no faster to recall negative personal memories,

but were considerably slower to retrieve positive personal memories. This result ought to alert us to the possibility that the increased probability of remembering depressing things from the past in emotionally disturbed patients may not be due to these memories being unusually fast, but due to rival positive memories being unusually slow (allowing the negative memories to win the accessibility race). Though the subjective experience is the same whether negative memories predominate because they are fast or because positive memories are slow, the therapeutic implications do differ depending on which explanation is correct. There would be little point in trying to slow down negative memories, but it would be important to speed up positive memories.

Using these specific assessment techniques we can also examine why this group of emotionally disturbed people took longer to retrieve positive memories. In two studies we have found that patients get stuck at the level of remembering general rather than specific memories. For example, patients quickly respond to the cue-word *safe* with the general memory "when I'm in bed," or to the cue-word *happy* with the general memory "the first few years of my marriage" but have difficulty getting beyond this to specific instances. It is a commonly occurring feature of therapy with depressed patients that they report a general memory about their past (e.g., that there were some good things in their life) but find it difficult to retrieve specific examples. Several researchers in the field of autobiographical memory have suggested that retrieval from long-term memory is a staged process in which a general description or context is first retrieved, then specified in more and more detail until a specific exemplar is retrieved (Norman & Bobrow, 1979; Reiser, Black, & Abelson, 1985). Evidence is accumulating that suggests that in our experiments mood affects the patients' memories by impairing these mechanisms of retrieval, blocking the production of memories at the point at which the system moves from accessing a general context to producing a specific example of that context. This blocking appears to occur more frequently for positive than negative memories, and it is this that largely accounts for the slower speed of recall of specific positive memories.

This is potentially important because up until now differential speed of recall of positive and negative memories has been explained mostly in terms of Anderson and Bower's associative network theory of memory (Anderson & Bower, 1973) in which emotion nodes spread activation to associatively linked nodes (see Chap. 11). If mood effects on memory are viewed within the network framework, remedial strategies have to concentrate on either changing the predominant mood or changing the patterns of associations of the network. But if mood disturbs retrieval strategies, as we have suggested, such therapies are likely to be incomplete. We shall need treatment strategies that directly address the deficit being shown, and this will require better descriptions of underlying retrieval processes than have been given hitherto.

CONCLUDING REMARKS

Cognitive behavior therapy seems to have established itself as a major supplement to drug therapy for unipolar nonpsychotic depression. Its use in psychotic depression (as defined by the presence of delusions or hallucinations) has not yet

been studied, so its value for these patients cannot yet be assessed. We still do not know how severe a patient has to be before he or she is inaccessible to cognitive procedures, though the fact that it consists of such a mixture of behavioral and cognitive techniques makes it more flexible than many psychotherapies. We do at least know now (Blackburn *et al.*, 1981; Murphy *et al.*, 1984) that low I.Q. and low social class is not a contraindication. We are also able to conclude that drugs and cognitive therapy may have similar effects on depressive cognitions, though we have seen that it is important to examine individual dysfunctional attitudes, and that, when this has been done (Reda, 1983) there are some negative attitudes that seem peculiarly resistant to pharmacological therapy. This had led us to consider the need for more specific psychological assessments of the processes purported to underlie negative thinking patterns. When this is done (e.g., with autobiographical memory tests) interesting data emerge concerning the distinction between, for example, general and specific memories.

Nevertheless, examining those studies that attempt to measure the processes underlying response to treatment, it is hard to escape the conclusion that the picture is very confused. Behavioral and cognitive components seem to be interchangeable in their effects (Zeiss *et al.*, 1979), yet mutually to complement each other (Taylor & Marshall, 1977). Different types of patients (e.g., endogenous and nonendogenous) respond differently to the same subcomponent (Fennell & Teasdale, 1984), yet respond the same way to the overall package (Beck, Hollan, Young, Bredosian, & Budenz, 1985; Blackburn *et al.*, 1981; Murphy *et al.*, 1984). Nonspecific treatment effects are of course a possibility but the most careful control of these by McLean and Hakstian (1979), using the same number of hours of treatment with relaxation and psychotherapy as was used in the cognitive behavior therapy, showed a clear superiority for the latter group.

What then of the mechanisms maintaining depression that cognitive and behavior therapy are affecting? Teasdale (1983) speculated that depression is maintained by the effects that small mood changes have on the recruitment of global negative self-evaluations. Many individuals may suffer identical mood shifts but they differ in the extent to which the same mood activates global evaluations. It is interesting to speculate that patients who possess a natural style that emphasizes self-initiated activity as a way of coping with dysphoric mood may have a learning history in which global negative evaluations have not taken such a grip, and in which increased activity "despite the mood" has been able to distance them from their pessimism. Individual differences in learning histories in these respects may be important predictors of response to treatment, and it is perhaps these that learned resourcefulness (Rosenbaum, 1980) and a favorable response to "Coping with Depression" (Fennell & Teasdale, 1987) reflect. What place then do more specific cognitive mechanisms have? If the speculations derived from Williams and Broadbent (1986b) are correct, the nature of memory is likely to be important. The tendency to recall general memories may prevent detailed information about events in the past from contributing to potentially useful strategies to cope with current problems. In other words, for any given learning history, there will be differences in how effectively it is brought to bear on current problems. Both the learning history and the memory for specific items in the history may prove to be crucial factors that the cognitive behavior

therapist needs to be making use of to change current behavior and perspectives and thereby reduce depression.

REFERENCES

- Abramson, L. T., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87*, 49–74.
- Anderson, J. R., & Bower, G. H. (1973). *Human associative memory*. Washington, DC: Winston.
- Beck, A. T. (1964). Thinking and depression: 2. Theory and therapy. *Archives of General Psychiatry, 10*, 561–571.
- Beck, A. T. (1967). *Depression: Clinical, experimental and theoretical aspects*. New York: Hoeber. (Republished as *Depression: Causes and treatment*. Philadelphia, PA: University of Pennsylvania Press, 1972).
- Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.
- Beck, A. T., Hollan, S. D., Young, J. E., Bedrosian, R. C., & Budenz, D. (1985). Treatment of depression with cognitive therapy and amitriptyline. *Archives of General Psychiatry, 42*, 142–148.
- Beck, A. T., Shaw, A. J., Rush, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York: Wiley.
- Becker, J. (1974). *Depression: Theory and research*. Washington, DC: V. H. Winston.
- Blackburn, I. M., Bishop, S., Glen, I. M., Whalley, L. J., & Christie, J. E. (1981). The efficacy of cognitive therapy in depression: A treatment trial using cognitive therapy and pharmacotherapy, each alone and in combination. *British Journal of Psychiatry, 139*, 181–189.
- Carney, M. W. P., Roth, M., & Garside, R. F. (1965). The diagnosis of depressive syndromes and the prediction of ECT response. *British Journal of Psychiatry, 111*, 659–674.
- Carver, C. S., & Ganellan, R. J. (1983). Depression and components of self-positiveness: High standards, self-criticism and overgeneralization. *Journal of Abnormal Psychology, 92*, 330–337.
- Chodoff, P. (1972). The depressive personality. *Archives of General Psychiatry, 27*, 666–673.
- Davis, E. (1982). *An investigation into the effect of internally and externally focused tasks on depressed mood*. Unpublished masters thesis, Dept. of Psychiatry, University of Newcastle-upon-Tyne.
- Fennell, M. J. V., & Teasdale, J. D. (1984). Effects of distraction on thinking and affect in depressed patients. *British Journal of Clinical Psychology, 23*, 65–66.
- Fennell, M. J. V., & Teasdale, J. D. (1987). Cognitive therapy of depression: individual differences in the process of change. *Cognitive Therapy and Research, 11*, 253–272.
- Godlib, I. H., & McCann, C. D. (1984). Construct accessibility and depression: An examination of cognitive and affective factors. *Journal of Personality and Social Psychology, 47*, 427–439.
- Hamilton, M. (1967). Development of rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology, 6*, 278–296.
- Harpin, R. E., Lieberman, R. P., Marks, I., Stern, R., & Bohannon, W. E. (1982). Cognitive behaviour therapy for chronically depressed patients: A controlled pilot study. *Journal of Nervous and Mental Disease, 170*, 295–300.
- Kihlstrom, J. E., & Nasby, W. (1981). Cognitive tasks in clinical assessment: An exercise in applied psychology. In P. C. Kendall, & S. D. Hollon (Eds.), *Assessment strategies for cognitive-behavioural interventions* (pp. 287–318). New York: Academic Press.
- Kuiper, N. A., & MacDonald, M. R. (1982). Self and other perception in mild depressives. *Social Cognition, 1*, 223–239.
- Lewinsohn, P. N. (1976). Activity schedules in treatment of depression. In J. D. Krumboltz & C. E. Thoresen (Eds.), *Counselling Method* (pp. 74–83). New York: Holtz, Rinehart, Winston.
- Lloyd, G. G., & Lishman, W. A. (1975). Effect of depression on the speed of recall of pleasant and unpleasant experiences. *Psychological Medicine, 5*, 173–180.
- Maier, S. F., & Seligman, M. E. P. (1976). Learned helplessness: Theory and evidence. *Journal of Experimental Psychology (General), 105*, 3–46.
- Mathews, A., & MacLeod, C. (1985). Selective processing of threat cues in anxiety states. *Behaviour Research and Therapy, 23*, 563–569.
- McLean, P. D., & Hakstian, A. R. (1979). Clinical depression: Comparative efficacy of outpatient treatments. *Journal of Consulting and Clinical Psychology, 47*, 818–836.

- Meichenbaum, D. (1977). *Cognitive behaviour modification*. Morristown, NJ: General Learning Press.
- Merluzzi, T. V., Rudy, T. E., & Glass, C. R. (1981). The information processing paradigm implications for clinical science. In T. V. Merluzzi, C. R. Glass, & M. Genest (Eds.), *Cognitive assessment* (pp. 77–126). New York: Guildford Press.
- Murphy, G. E., Simons, K. D., Wetzel, R. D., & Lustman, P. J. (1984). Cognitive therapy and pharmacotherapy; singly and together in the treatment of depression. *Archives of General Psychiatry*, *41*, 33–41.
- Norman, D. A., & Bobrow, D. G. (1979). Descriptions: An intermediate stage in memory retrieval. *Cognitive Psychology*, *11*, 107–123.
- Phillips, L. W. (1981). Roots and branches of behavioural and cognitive practice. *Journal of Behaviour Therapy and Experimental Psychiatry*, *12*, 5–17.
- Reda, M. (1983). Cognitive organization and antidepressants. In M. A. Reda & M. J. Mahoney (Eds.), *Cognitive psychotherapies* (pp. 119–140). Cambridge, MA: Ballinger.
- Reiser, B. J., Black, J. B., & Abelson, R. P. (1985). Knowledge structures in the organization and retrieval of autobiographical memories. *Cognitive Psychology*, *17*, 89–137.
- Rosenbaum, M. (1980). A schedule for assessing self control behaviours: Preliminary findings. *Behaviour Therapy*, *11*, 109–121.
- Rotzer, F. T., Nabit, U., Koch, H., & Pflug, B. (1982). *Zur Bedeutung von Attribuierungsprozessen bei der Depressionsbehandlung*. In *Leur G.* (Ed.) *Bericht des 33 Kongresses der deutschen Gesellschaft für Psychologie*. September 1982. Mainz, FRG.
- Rush, A. J., Beck, A. T., Kovacs, M., & Hollon, S. (1977). Comparative efficacy of cognitive therapy and pharmacotherapy in the treatment of depressed outpatients. *Cognitive Therapy and Research*, *1*, 17–37.
- Seligman, M. E. P. (1975). *Helplessness: On Depression, development and death*. San Francisco: W. H. Freeman.
- Simons, A. D., Garfield, S. L., & Murphy, G. E. (1984). The process of change in cognitive therapy and pharmacotherapy for depression. *Archives of General Psychiatry*, *41*, 45–51.
- Spitzer, R. L., Endicott, J., & Robins, E. (1978). *Research Diagnostic Criteria (ROC) for selected group of functional disorders* (3rd ed.), N.Y. State Psychiatric Institute, Biometrics Research, New York.
- Strack, F., Schwarz, N., & Gschneidinger, E. (1985). Happiness and reminiscing: The role of time perspective, affect, and mode of thinking. *Journal of Personality and Social Psychology*, *49*, 1460–1469.
- Taylor, F. G., & Marshall, W. L. (1977). Experimental analysis of a cognitive or behavioural therapy for depression. *Cognitive Therapy and Research*, *1*, 59–72.
- Teasdale, J. D. (1983). Negative thinking in depression: Cause, effect or reciprocal relationship? *Advances in Behaviour Research and Therapy*, *5*, 3–25.
- Teasdale, J. D., Fennell, M. J. V., Hibbert, G. A., & Amies, P. L. (1984). Cognitive therapy for major depressive disorder in primary care. *British Journal of Psychiatry*, *144*, 400–406.
- Teasdale, J. D., & Fogarty, S. J. (1979). Differential effects of induced mood on retrieval of pleasant and unpleasant events from episodic memory. *Journal of Abnormal Psychology*, *88*, 248–257.
- Teasdale, J. D., & Resin, V. (1978). The effects of reducing frequency of negative thoughts on the mood of depressed patients—Tests of a cognitive model of depression. *British Journal of Social and Clinical Psychology*, *17*, 65–74.
- Watts, F. N., & Sharrock, R. (1985). Description and measurement of concentration problems in depressed patients. *Psychological Medicine*, *15*, 317–326.
- Watts, F. N., McKenna, F. P., Sharrock, R., & Trezise, L. (1986). The stroop test as a measure of phobic attention. *British Journal of Psychology*, *77*, 97–108.
- Williams, J. M. G. (1984a). *The psychological treatment of depression: A guide to the theory and practice of cognitive-behaviour therapy*. London & Canberra: Croom Helm.
- Williams, J. M. G. (1984b). Cognitive behavior therapy for depression: Problems and perspectives. *British Journal of Psychiatry*, *145*, 254–262.
- Williams, J. M. G., & Broadbent, K. (1986a). Distraction by emotional stimuli: use of a stroop task with suicide attempters. *British Journal of Clinical Psychology*, *25*, 101–110.
- Williams, J. M. G., & Broadbent, K. (1986b). Autobiographical memory in suicide attempters. *Journal of Abnormal Psychology*, *95*, 144–149.
- Williams, J. M. G., & Nulty, D. (1986c). Construct accessibility, depression and the emotional stroop task: transient state or stable structure. *Personality and Individual Differences*, *7*, 485–491.

- Wilson, P. H., Goldin, J. C., & Charbonneau-Powis, M. (1983). Comparative efficacy of behaviour and cognitive treatments of depression. *Cognitive Therapy and Research*, 7, 111-124.
- Zeiss, A. M., Lewinsohn, P. M., & Munoz, R. F. (1979). Non-specific improvement effects in depression using interpersonal skills training, pleasant activity schedules, or cognitive training. *Journal of Consulting and Clinical Psychology*, 47, 427-439.

CHAPTER 13

Cognitive Theories of Motivation

Chris R. Brewin

INTRODUCTION

The theoretical bases of behavior therapy have traditionally been concerned with the acquisition and extinction of fear and avoidance responses, and have drawn extensively on animal learning experiments carried out under carefully controlled conditions. The intriguing parallels with human experience to be found in animal experiments should not, however, be allowed to detract from an appreciation of the differences between the laboratory and the situation of a patient in a health care delivery system. In the latter, for instance, patients have already selected themselves as being unable to solve their problems on their own, and have defined these problems as requiring the sort of help provided by professionals. Patients also vary considerably in their perceptions of themselves and their abilities, and many complain of chronically low self-esteem. These factors are likely to interact with response to treatment, and for these reasons such individuals may not provide typical examples of fear acquisition and extinction processes of the kind that would be common in the laboratory.

Another important factor in generalizing from the laboratory is that human beings have very large behavioral repertoires that give them many options in responding to unpleasant emotional states or unwanted behaviors. They are likely to have exercised or still be exercising some of these options when they consult a behavior therapist, and some of these solutions may themselves be barriers to therapeutic progress. The therapist must present such a convincing and acceptable analysis of the patients' problems that they abandon competing options and devote maximum effort to following the therapist's suggestions. In practice, of course, many patients remain unconvinced by the analysis offered, make half-hearted attempts to carry out the suggested plan, and terminate treatment prematurely. Although the incidence of failure to comply with psychological treatments has not been so extensively documented as is the case with physical treatments (e.g., Johnson, 1981; Ley, 1977), the problem of resistance is coming to be accepted as of major importance (Wachtel,

1982). It is essential that the theoretical base of behavior therapy recognizes and includes an account of why certain response options are selected and why there are such wide variations in the vigor, persistence, and effectiveness with which people attempt to deal with their symptoms.

The practice of behavior therapy and cognitive-behavioral therapy has largely been based on the assumption that certain people need to be taught effective techniques for controlling symptoms. According to this skills model, what distinguishes patients who require treatment is that they lack the knowledge to deal with their symptoms effectively. Therapy is mainly a question of analyzing the functional relations between person, environment, and symptoms, and teaching the appropriate procedures. An alternative view is that patients know (or at least suspect) what they ought to do, but for one reason or another abandon their attempts to deal with their symptoms before they have a chance to be effective. In other words, however cooperative the patient, it is possible to see the problem that has brought them to the clinic not only in terms of lack of knowledge but also of lack of motivation.

According to this motivational perspective, behavior therapy does not so much educate people as provide them with the encouragement and incentive to persist with promising coping strategies. Its function is to support coping attempts and prevent or counteract the development of hopelessness and helplessness by enhancing perceptions of control over symptoms and over the environment more generally. At present the evidence for this point of view is only indirect, but comes from at least three sources. First, the techniques of behavior therapy, such as the location of reinforcing events, the breaking down of tasks into graded hierarchies, and the encouragement and support of the therapist, all suggest the importance of motivational processes over and above the simple provision of instruction. Second, patients do tend to feel generally powerless and to believe that they have less control over their lives than do nonpatients, perceptions that change during the course of therapy in the direction of greater control (Strickland, 1978). There are indications that beliefs about control also predict the length of time people take to recover from disorders such as depression (Brewin, 1985). The third line of evidence comes from studies that have manipulated perceptions of control, for instance by persuading people to attribute their improvement to an internal factor such as effort or to an external factor such as a drug. The results indicate that self-attribution of improvement at the end of treatment is associated with superior outcome at follow-up (see Brewin & Antaki, 1982, 1987, for reviews).

Cognitive theories of motivation have been dominated by a very simple idea, namely that the intensity or persistence of behavior is determined by a combination of the value of the goal the person is trying to achieve and the expectancy that some behavior will be effective in attaining that goal. Little effort will be expended when the goal is unimportant or when the behavior is seen as unlikely to achieve the desired ends. Sometimes the goal is very attractive, such as overcoming life-long shyness, but the expectancy that one will change is too small to encourage social experimentation. Similarly, little effort might be exerted to overcome a tic or nervous habit because, although confident of success, the individual did not consider it sufficiently important to compensate for the stigma of referral to a psychologist or psychiatrist.

These ideas, which are usually known as expectancy/value theory, have a long history and can be found in the work of psychologists such as Tolman, Lewin, and Atkinson. They also underpin many of the most influential theories applied to the clinical domain, such as Rotter's social learning theory, Bandura's social learning theory, and Seligman's learned helplessness theory. Although not ignoring the value component, these three have all placed particular emphasis on the importance of expectancies, Rotter by articulating the distinction between generalized and situation-specific expectancies, Bandura by emphasizing the distinction between outcome and efficacy expectancies, and Seligman by placing the expectation of uncontrollability at the heart of his explanation of cognitive and motivational deficits in behavior. Before examining the role of expectancy and value judgments in more depth, however, it is necessary to address the issue of what is meant by a cognition and how cognitive processes are related to behavior.

TWO COGNITIVE SYSTEMS

Recent approaches to psychological therapy, such as those of Beck and Ellis, have been described as cognitive for a number of reasons. First, they emphasize the importance of paying attention to patients' reported attitudes and beliefs, and assume that changing these beliefs can alleviate disturbed mood. However, the cognition in cognitive therapy does not refer solely to thoughts that the patient can report. Both Beck and Ellis have emphasized that patients may not be aware of the rules that guide their behavior, and that may only become explicit as a result of therapy. For instance, Beck, Rush, Shaw, and Emery (1979) hypothesized the presence of core assumptions, unarticulated rules by which the person tries to integrate and assign value to the raw data of experience. Beck *et al.* have also focused on the occurrence of automatic thoughts in depressed patients, thoughts that occur so rapidly and spontaneously in certain situations that patients are often barely aware of their existence.

Cognition and conscious thought are, of course, quite different concepts. The term *cognition*, although it may be used in a number of ways, commonly refers to the representation within an organism of knowledge about itself or its environment (e.g., Mandler, 1984). This knowledge, for instance concerning covariation between stimuli, behavior, and outcomes, is essential to inform the organism about the likely consequences of its actions and about the rewards and punishments it might anticipate in different situations. The presence of cognitive processes for integrating information is inferred from observations of behavior under different circumstances, and it appears that the processes by which covariation is assessed are remarkably similar in animals and human beings (Alloy & Tabachnik, 1984). In contrast, conscious thought describes a phenomenological experience of awareness that is restricted by the inability of human beings to attend to more than a few events or experiences at a time. Neisser (1967) argued that people do not have access to cognitive processes themselves, but only to the products of those processes, and that the vast majority of cognitive acts

and processes take place out of awareness. Conscious thought acts to select and elaborate cognitions that are otherwise present only fleetingly, if at all, in awareness.

In parallel to the distinction between these two cognitive systems, a number of psychologists have documented the presence of two forms of information processing, one under the control and one not under the control of the individual (Bargh, 1984; Posner & Snyder, 1975; Schneider & Shiffrin, 1977). Conscious or control processes have been described as highly flexible and adaptable but with a processing capacity severely limited by the attention span of the human being. This limited capacity means that judgments made under conditions of uncertainty or complexity are strongly influenced by heuristics (e.g., representativeness and availability; Tversky & Kahneman, 1974) or by *a priori* causal hypotheses. Conscious processing is deliberate and effortful, and can be easily disrupted when there are other stimuli competing for the individual's attention. Automatic processes, on the other hand, are rapid, require minimal attention to occur, and may be activated without intention or awareness. In attributing causality, for instance, people are powerfully influenced by the fact that two events occur close together in time. They are also influenced by the salience of stimuli, preferring to see more salient stimuli as causally prior. These factors can affect attributions without the knowledge or awareness of the individual (Kassin & Baron, 1985). Automatic processes are also likely to be activated in situations with which the individual is very familiar, whereas novel situations are more likely to elicit the conscious, deliberate type of processing.

For clinical psychologists two very important questions arise from the distinction between these two systems or forms of processing. First, How good or bad is our potential conscious access to our mental states? In other words, How can we consciously know what we feel or expect if these cognitions are the product of largely automatic or nonconscious processing? One view is that some initially nonconscious cognitive material can become accessible to consciousness without any great distortion happening in the process. This assumption is common to many forms of psychodynamic therapy, such as those that aim for the recovery of forgotten memories together with their associated emotions. Beck *et al.* (1979) also assume that core assumptions of which the patient is not aware can be identified from the content of automatic thoughts, the types of thinking errors made, and the frequent use of particular words. These underlying rules must be carefully inferred on the basis of observational data, but Beck *et al.* do suggest that introspection by the patient is useful and that the aim is for the therapist to articulate assumptions that will be recognized by the patient as corresponding to previously inaccessible dysfunctional beliefs.

In contrast, other social and cognitive psychologists have suggested that when people describe their feelings or explain their behavior they are simply making inferences designed to account for or justify their actions. This recent scepticism derives from a large body of research showing that people often have poor access to their mental states (Nisbett & Wilson, 1977; Wilson, 1985). Verbal reports, in other words, constitute guesses about the appropriate labels for certain internal states ("feelings"), or about the real, underlying causes of behavior. Under certain circumstances, for instance when the relevant stimuli are readily observable, when the cause-effect

relations correspond to *a priori* beliefs, and when there is a premium on accurate assessment, such guesses may correspond well to actual environmental contingencies. Under other circumstances, for instance when there are a large number of relevant stimuli, when the actual cause–effect relations are unexpected, and when there are other influences, such as the desire to preserve a good image in one’s own eyes or those of others, guesses are unlikely to be accurate reflections of environmental contingencies.

The second, and related, question concerns the mediation of behavior. For many years behaviorists have argued that behavior is shaped by environmental events through processes such as operant and classical conditioning, and that conscious beliefs and feelings are by-products of these processes rather than causal influences in their own right. These assertions appear to be supported by the experiences of many psychiatric patients. The phobic or obsessional patient may respond with fearful sensations, images, and thoughts to certain situations, but without knowing why he or she does so and in spite of the conscious belief that there is no objective reason to be frightened. Similarly, individuals may find themselves responding angrily or irrationally to employers, colleagues, marital partners, and sometimes their therapists, without being aware of any obvious provocation.

Social psychological research has provided many other examples. Latané & Darley (1970) found, for example, that the probability of a person helping another in distress is related to the number of other people present, but that individuals asked to explain their behavior denied that they had been or would be affected by this factor. Nisbett & Wilson (1977) reported a number of findings indicating that people are frequently unaware of the influences affecting their behavior, and concluded that to ask people about their cognitive processes might be highly misleading. Additional evidence comes from investigations of the effect of examining one’s mental state on attitude–behavior consistency. If introspection does involve direct access to one’s mental states, then the consistency between one’s reported feelings and actual behavior should increase with greater introspection. Wilson, Dunn, Bybee, Hyman, and Rotondo (1984) tested this hypothesis in a number of situations, including getting couples to rate how happy they were with their relationship and then finding out whether they were still dating several months later. The more couples were encouraged to introspect before rating their relationship, the lower was the correlation between these ratings and subsequent outcome. Similar results were obtained in the other situations, indicating that self-analysis tended to produce greater error, not greater accuracy.

There is, then, ample evidence that conscious thought is not a necessary determinant of behavior. This fact does not, of course, say anything about whether conscious thought may (sometimes) determine behavior, and under what conditions, or whether conscious thought merely predicts behavior on those occasions in which its content is an accurate reflection of the content of underlying, nonconscious cognitions. To begin with, it is appropriate to ask whether there are any circumstances in which verbal reports might be expected to be accurate. Morris (1981) has noted that Nisbett and Wilson’s (1977) examples detailing the inaccessibility of mental states involve situations in which people have to account for their behavior rather than report what

they intend to do next, and he argues that there is a fundamental difference between these two types of situation. To ask people whether they are going to go for a walk (an intended action) involves a different kind of verbal report from asking them why they liked someone or whether they are going to like someone (nonintended actions). Morris accepts that Nisbett and Wilson have cast doubt on people's ability to explain nonintended behaviors, but maintains that there is no reason to extend the same skepticism to people's reports of their intentions.

The distinction between intended and nonintended, or regulated and unregulated, actions is indeed a vital one for interpreting the results of research in this area. Whereas self-reports are usually regulated by the person, behavior may be regulated or unregulated. For instance, in social interaction facial expression is more often regulated than body cues or tone of voice (Ekman & Friesen, 1974; Zuckerman, Larrance, Spiegel, & Klorman, 1981). Wilson (1985) argued that a person's true feelings or beliefs are demonstrated through these unregulated channels, whereas regulated behaviors indicate what a person is consciously trying to convey. The accuracy of beliefs about underlying mental states will therefore be estimated best by the correspondence between such beliefs and unregulated behaviors.

To expand on this line of argument, there should generally be good agreement between verbal reports, particularly those involving short-term intentions, and regulated behavior, because both are under the person's control and can relatively easily be made consistent. This agreement should weaken as the intentional elements in the report become less prominent or specific, or as the need for consistency between verbal report and behavior diminishes. Verbal reports will only agree with unregulated behavior to the extent that they represent relatively accurate guesses at underlying mental states. In support of these ideas Wilson, Lassiter, and Stone (1984) found that subjects' reported preferences for moderately versus extremely similar others predicted their regulated social behaviors (facial expression, body inclination, and talking time) in the presence of a preferred versus nonpreferred other, but did not predict the subjects' unregulated social behaviors (interpersonal distance, eye contact, and body orientation).

To summarize, the value of self-reports would appear to lie more in their relation to intentional future actions than in any insight they might provide into complex feeling states or into the contingencies governing past behavior. At times verbal reports will correctly identify complex mental states or cause-effect relations, particularly when the causes are salient ones and the causal link conforms to *a priori* beliefs. It is important to remember, however, that such reports are based mainly on inferences rather than on direct access to internal states, although the possibility of a degree of access to these states in a more prolonged therapeutic context cannot be ruled out. Whether these reports are accurate or not, however, conscious self-referent thought is likely to determine a person's intentional actions. In parallel to this system, cognitions not available to consciousness have the capacity to exert their own influence on behavior. In the remainder of this chapter I discuss how the operation of these two cognitive systems may be further specified in terms of expectancies and values, and how they may influence response to physical and psychological disorders. In particular, appraisal of illness or other adversity is likely to influence response to physical and psychological disorders.

EXPECTANCIES

Expectancies as determinants of behavior have been studied and analyzed in considerable detail. In Rotter's (1954) social learning theory, for instance, there are two main kinds of expectancy, generalized and situation specific. Briefly, Rotter says that when we assess how successful our actions are going to be we utilize not only our knowledge of that particular situation but also beliefs that have developed through the lifespan concerning more general capabilities, beliefs that summarize a wide range of disparate experiences. One particular example of a generalized expectancy that has probably attracted a disproportionate amount of research is locus of control or, to give it its full title, generalized expectancies for internal-external control of reinforcement. This construct reflects the degree to which individuals believe that reinforcements, that is, the good and bad things that happen to them, are under their own control or under the control of external factors, such as chance. Rotter suggests that this kind of generalized expectancy will be more likely to predict people's behavior when they are in novel or unfamiliar situations, whereas in familiar situations specific expectancies are more likely to be appropriate.

Strickland (1978) has provided a useful review of studies that have related self-report measures of locus of control to the attitudes and behavior of people with physical and psychological disorders. The more anxious or depressed a person is, the more external their locus of control tends to be. Furthermore, people usually become more internal as treatment progresses. There is also some evidence that internals respond better to treatments that give them greater control and responsibility, whereas externals prefer treatments with greater structure. Strickland concludes her review as follows:

Although results are not altogether as clear, convincing, and as free from conflict as one might hope, the bulk of the research is consistent in implying that when faced with health problems, internal individuals do appear to engage in more generally adaptive responses than do externals. These range from engagement in preventive and precautionary health measures through appropriate remedial strategies when disease or disorder occurs. (p. 1205)

Many criticisms of measures of generalized expectancies such as locus of control (e.g., Brewin & Shapiro, 1984) have pointed to the failure to distinguish various more specific beliefs. In one sense this criticism is unfair because the scales were designed to tap general expectations that might be expected to influence behavior when people found themselves in unfamiliar situations. As we have seen, locus of control is related to a very wide range of behaviors, and criticisms should perhaps not be leveled at the scale so much as at the researchers who have not used it appropriately. But the generality of the scales does mean that subjects' scores may reflect a variety of quite different beliefs, and that results using it are therefore hard to interpret. For these reasons, and in the interests of further improving the prediction of behavior, it has been found necessary to develop expectancy measures that focus on specific situations with which patients may be familiar and on specific beliefs.

Bandura (1977) reminded us that two classes of specific expectancies can be distinguished, outcome expectancies and efficacy expectancies. The former are concerned with people's beliefs about the likely success of a treatment or other course of action, and the latter with their beliefs about their own ability to actually carry

out those actions. Thus it is possible to imagine a person who believes flooding to be an effective form of treatment for his phobia (high outcome expectancy) but does not believe himself capable of exposing himself to his most feared situation (low efficacy expectancy).

The importance of beliefs about the effectiveness of treatments (outcome expectancies) is amply demonstrated by one of the most reliable psychological phenomena in medicine, the placebo effect. A placebo is traditionally a pharmacologically inert substance that a patient takes in the belief that it is an active drug. When such a substance leads to changes in the patient's state it is said to produce a placebo effect. These effects are so widespread that placebo controls are routinely included in tests of drug action: a drug is four to five times more likely to be reported as effective if there is no placebo control than if there is (Shapiro, 1971). Placebos have been shown to have a beneficial impact on a huge variety of disorders, including dental pain, asthma, multiple sclerosis, the common cold, diabetes, ulcers, and Parkinson's disease. Placebos have also been known to be so convincing that their termination produced withdrawal symptoms.

Outcome expectancies have also been considered important in the design and interpretation of research on psychological therapies, and it has been common to compare treatments such as systematic desensitization with supposedly ineffective placebo treatments to control for these expectations. Unlike drugs, however, whose efficacy cannot be readily assessed by the patient, psychological therapies can be assessed in this way and patients can form views about which are most likely to be effective. Many placebo manipulations have been found to be intrinsically less credible than the real treatments with which they were being compared, thus biasing studies in favor of the latter. When treatments such as systematic desensitization are compared with equally credible control procedures, their superiority is reduced and often disappears altogether (Lick & Bootzin, 1975). Among the most plausible reasons for this Lick and Bootzin suggested that high expectancies might encourage subjects to test the hypothesis "I am cured" by exposing themselves to real-life phobic stimuli. Other possibilities were that subjects might be responding to the demand characteristics of the situation, or that high expectancies might produce reassuring self-talk (such as "I'm really not afraid of heights, I don't have anything to worry about") which would directly reduce anxiety.

Bandura (1977) labeled expectations about one's own performance self-efficacy beliefs. His thesis is that people vary greatly in the confidence that they feel when coping with difficult situations, particularly ones that produce unpleasant emotional arousal. The more confident they feel about responding skillfully to the varying demands of the situation or, in his terms, the greater their self-efficacy, the harder they will try to overcome the problem and the longer they will persist at it. From this Bandura deduces that successful psychological therapies are those that are most effective at increasing self-efficacy: it is a mechanism that accounts for why some therapies are generally better than others and why some people do better than others at the same therapy.

A number of investigations into the relation between self-efficacy and avoidance behavior have been conducted, and these are described in Bandura (1977, 1982). Different treatments appear to have widely varying effects on subjects' self-efficacy

ratings. As one would expect, performance-based treatments increase efficacy more than imaginal treatments. Whatever the treatment, however, there is a close relation between the magnitude and strength of efficacy judgments and subsequent performance. Self-efficacy has now been found to be related to a range of behaviors in several different situations. It predicts the performance of agoraphobics carrying out tasks they find difficult, such as walking alone, shopping, and dining in a restaurant (Bandura, Adams, Hardy, & Howells, 1980). It also predicts the behavior of socially anxious males (Barrios, 1983), assertiveness (Kazdin, 1979), achievement behavior (Bandura & Schunk, 1981), and self-control of addictive behaviors such as smoking (Conditte & Lichtenstein, 1981; DiClemente, 1981). Self-efficacy therefore appears to be a construct that is not just relevant to the treatment of fear but that has a general motivational role.

Many of the criticisms of self-efficacy theory seem to have missed this point, that it is a general theory of motivation and not one of fear acquisition or reduction. Bandura is not really concerned with why a person originally became afraid of snakes nor with the ideal therapy for phobias. What he is concerned about is why people do not cope equally successfully with their fears, with their marital problems, with their careers, with their poor tennis service, or with anything else that constitutes a problem. It is therefore true, but irrelevant, to argue that Bandura places too little weight on the acquisition of conditioned emotional reactions or that exposure rather than self-efficacy is the mechanism of fear reduction. Therapy outcome depends not only on knowing the correct treatment but on persistently and determinedly performing it.

The predictive power of self-efficacy is good in studies of snake phobics and smokers, but not all studies have got such good results (Lane & Borkovec, 1984; Meier, McCarthy, & Schmeck, 1984). These findings suggest an important limitation to the theory. Predictive power appears to be best when (a) the criteria of a successful performance are unambiguous, and (b) the behavior necessary is under deliberate conscious control. The snake phobics in Bandura's studies were aware of precisely what each test consisted of, and could decide whether to do it or not. Similarly, the smokers had a conscious decision whether or not to abstain. Controlling one's non-verbal behavior or writing a good essay are by contrast tasks over which one has far less control. One may intend to perform well but by the nature of the task get little feedback about whether one is succeeding. Under these circumstances it may be difficult to recognize when one's performance is failing to live up to one's expectations.

These limitations in their turn suggest two things. First, efficacy judgments may be no more than a statement of what a person intends to do and, second, people may only be able to make valid efficacy judgments under certain conditions. Obtaining self-efficacy judgments from smokers or phobics about to take a behavioral avoidance test appears very similar to measuring their intentions. Does the measurement of efficacy offer any advantage over simply asking people what they are going to do next? Although it is indeed difficult to distinguish between efficacy judgments and statements of intention when the action is very shortly to follow, the distinction does become useful when people fail to do what they intend. We can then ask whether efficacy was higher in those who fulfilled their intentions than in those who did not fulfill them. A recent study by DiClemente, Prochaska, and Gibertini (1985) indicates

that people who fulfill their intention to give up smoking or to remain abstinent do have greater prior levels of efficacy than those who fail to carry through their plans. This means that to know about a person's level of efficacy is more informative than simply to know about their intention (at least over longer periods).

The second point has to do with the criteria for accurate introspection. Bandura implies by his use of the term *self-referent thought* that self-efficacy is exclusively a conscious process that people introspect about when they are asked to make efficacy ratings. In any case, the ratings certainly reflect conscious thought processes. Presumably a person can much more easily make a judgment about whether they will perform a simple voluntary action, on its own and with no distractions, than they can predict the level of a complex performance made up of many constituent parts that they are not in the habit of monitoring. This is similar to Ericsson and Simon's (1980) point that accurate introspection or verbalization has to be based on information in either short-term or long-term memory. In the same way, efficacy judgments will be inaccurate predictors of behavior that cannot be fully attended to. As Bandura notes, they will also be inaccurate predictors of the performance of skills that have only partially been mastered.

Another way of looking at this is to consider the sorts of experiences that might influence dealing with a feared situation. The fact that people often do not know why they are frightened of something means that they may not be aware of many of these experiences, perhaps because they occurred in childhood. Self-efficacy as a conscious thought process must then be limited by people's access to those experiences that are indeed relevant. Psychological treatments provide them with new information of a more or less compelling kind, but the information still has to be appraised in the light of preexisting information whose nature is largely unknown and that may be in conflict with the evidence provided by the therapy. Bandura assumes that the products of this appraisal process will be available to consciousness but this assumption is a dubious one. It is just as likely that this preexisting information, for instance in the form of learned associations between certain actions, rewards, and punishments, will exert its own influence on behavior outside of conscious awareness, although people will also be prompted by their conscious self-efficacy judgments to expend more or less effort on coping behavior. Efficacy is not so likely to be a good predictor of behavior when there are strong contrary expectations outside of awareness and when the nature of the behavior is such that it is difficult to achieve simply by regulating effort expenditure. Much of behavior cannot be regulated by deliberately trying harder and therefore will not reflect so readily the influence of self-generated motivation.

VALUES AND GOALS

The other general determinant of the direction, intensity, and persistence of behavior is value. This term refers to people's preferences for certain goals over others, and may be expressed either behaviorally by people's choice of activities or verbally in self-reports. Although values may derive from several sources, including innate

approach–avoidance tendencies, we shall be particularly concerned with learned preferences that arise through cultural transmission or the individual's unique experience. In either case it is assumed that values are adopted or rejected because of their previous associations with reward and punishment, either experienced or observed. Thus values such as caring for old people or avoiding causing pain might be deliberately inculcated by parents or teachers, or acquired through observation of admired others.

In practice, people have hierarchies of goals relating to different areas of their lives. On the most general level, people usually wish to maximize pleasant experiences and minimize unpleasant experiences such as physical pain, fear, anxiety, guilt, and cognitive inconsistency or dissonance. In order to do this successfully they strive to gain accurate information about their own abilities and characteristics (Festinger, 1954), and to understand, predict, and control their environment (Heider, 1958; Kelley, 1967). More specifically, they may wish to have an interesting job, an involving family life, intelligent and well brought-up children, financial security, and a certain amount of excitement and unpredictability. These goals must all coexist with others relating to moral values and standards of conduct, social and family obligations, and valued activities of purely personal interest, such as sport or hobbies.

It is evident that not all these goals can be pursued simultaneously, and that a system of priorities must be operated in order to resolve conflict. In certain situations it may be useful to know what people's goal priorities are in order to predict their reactions. For instance, Beck (1983) and Blatt, Quinlan, Chevron, McDonald, and Zuroff (1982) have proposed that depressives can be classified according to whether their goals are primarily concerned with achievement (self-critical or autonomous type) or whether they concern relationships with other people (dependent or sociotropic type). There is evidence that the latter are more likely to be depressed by a life event affecting their relationships than one affecting their work, whereas the former group show the opposite pattern of vulnerability (Hammen, Marks, Mayol, & DeMayo, 1985).

Many people have suggested that, other things being equal, priorities are organized in relation to a superordinate goal, the maintenance of self-esteem (e.g., Wylie, 1979). According to this view, people are motivated to preserve a positive image of themselves, both in their own eyes and in those of other people. One way in which they do so is through the use of self-serving biases (Bradley, 1978), the tendency to take credit for success and to attribute blame for failure externally. At times the goal of self-esteem maintenance may conflict with the goal of symptom reduction, a possibility that has been explored in research on self-handicapping strategies. Following Alfred Adler, Jones and Berglas (1978) suggested that symptoms might have value to a person faced with a threat to his or her self-esteem, because they would provide a potential alibi for failure whereas any success would be enhanced by the knowledge that it was achieved despite a handicap. As Snyder and Smith (1982) put it:

By adopting a symptom, the person increases the ambiguity as to the "real" underlying reason for a possible failure. . . . Behind the veil of ambiguity supposedly generated by the symptoms, the self-handicapper therefore is able to nurture a fantasy of self-esteem and competence.

Smith, Snyder, and Perkins (1983) argued that hypochondriacal individuals may use the symptoms of physical illness as a self-handicapping strategy. They selected female students who scored either high or low on the MMPI Hypochondriasis scale and told them that they would be taking a test of social intelligence. Before completing the second part of this test, which they expected to be difficult, subjects were either instructed that symptoms of physical illness would not affect their performance, or were given no instructions about the relation between symptoms and performance. Subjects then completed a questionnaire about their health in the past year and over the previous 24 hours. Hypochondriacal subjects in the illness-has-no-effect condition reported fewer health problems than those in the no-instructions condition, whereas this manipulation had no effect on nonhypochondriacal subjects. These results are consistent with the use of symptom report as a self-handicapping strategy by hypochondriacal individuals.

Many symptoms and behaviors, such as shyness, depression, test anxiety, and drug and alcohol abuse, may be viewed in these functional terms. Of particular clinical importance is the possibility that effort withdrawal, or failure to cooperate fully with treatment, reflects a similar strategy. Although there are no studies investigating this possibility with clinical samples, a number of very suggestive findings have been reported. For instance, it has been found that students who expected to fail an achievement test reduced their intended effort, but only when the task was an ego-involving one (Pyszczynski & Greenberg, 1983). In another study, high and low self-handicapping members of a men's swimming team were found to practice a similar amount prior to an unimportant meet, but high self-handicappers practiced much less than their low self-handicapping colleagues prior to an important meet (Rhodewalt, Saltzman, & Wittmer, in press). An interesting corollary is that self-handicapping may be avoided, and performance enhanced, by providing people with a ready-made explanation for failure that does not cast doubt on their abilities. After failure on an insoluble task, Frankel and Snyder (1978) found that subjects worked harder and were more successful at solving anagrams described as extremely difficult than they were at anagrams described as moderately difficult. This effect is contrary to what one might intuitively expect and indicates effort is not only affected by expectancy of success but also by the anticipated consequences of failure.

So far we have considered some of the values that may determine the general direction of a person's behavior, whether or not the person is consciously aware of them. In addition it is important to recognize people's capacity for deliberately and consciously regulating their own behavior and the amount of effort they choose to expend on attaining their goals. Kanfer (1970) proposed that when people's behavior is interrupted or fails to produce the intended effects, a process of self-regulation is initiated. This has three stages, of which the first is self-observation. People begin to attend to their behavior and try to reconstruct or monitor their actions. The second stage is self-evaluation, a comparison between these actions and internal standards of what the actions ought to have been. Any discrepancy can then be corrected, at least in theory. The final stage is self-reinforcement, when people either reward themselves for achieving their standards or punish themselves for failing to achieve them.

A similar view of self-regulation is put forward by Bandura (1978). Bandura notes that the stage of self-evaluation includes a number of components, such as the comparison of one's performance with personal and social norms, and the attribution of the cause of one's performance to personal and external factors. Faced with the same objective outcome, there is thus scope for enormous differences in evaluations, which will depend *inter alia* on past experience in similar situations, the choice of others with whom to compare oneself, and predispositions to make certain kinds of attribution. The more performance falls short of what is expected, and the more this is attributed to internal factors, the more people tend to strive to make up the deficit. This effect has been illustrated by research on industrial accident victims (Brewin, 1984). Victims who perceived themselves to be more negligent and to blame for their accidents returned to work significantly faster than their counterparts who did not feel at fault.

Although these cognitive activities could take place out of conscious awareness, Bandura points out that it is open to people to deliberately set themselves performance standards that they try to achieve, thereby guiding and regulating their efforts. Provided the activities can be voluntarily controlled, higher standards will lead to greater effort expenditure. When voluntary control is not so certain, setting subgoals that are relatively easily attainable is likely to sustain self-motivation (Bandura & Simon, 1977). People can also try to influence their own behavior by creating self-incentives and rewarding or punishing themselves contingent on their performance. Although these rewards and punishments will usually take the form of feelings of pride or blame, a number of studies have demonstrated the beneficial effects on performance of giving oneself tangible rewards, which can be as effective or more effective than externally arranged incentives.

CLINICAL IMPLICATIONS

Using an expectancy/value framework I have outlined a number of possible mechanisms that might account for patients' lack of involvement in therapy. The first involves low expectations of success. These may on occasion be pervasive and reflect a view of the self as generally ineffective or incompetent. In other patients a general feeling of competence may coexist with specific low outcome or efficacy expectations directed either at the treatment offered or at the behavior required of them. Other mechanisms concern the existence of competing values or goals that may not have been made explicit. In particular therapists should carefully consider the implications for the patient's self-esteem of attending therapy sessions, failing therapeutic objectives, and having to give up potential alibis for lack of success elsewhere in life.

It has also been argued that behavior is under the control of at least two systems that are in some sense cognitive. On the one hand behavior is influenced directly by cognitive representations of past experiences, which contain a record of stimuli associated with reward and punishment and information about the contingent relations between different types or classes of event. These representations are continually

being updated by the addition of new information gleaned from experience, observation, books, conversations, etc. In any event, people's access to these representations is highly restricted and they must infer the contents of the representations from their own behavior. Their conscious inferences and wishes may in turn exert their own influence on behavior via the self-regulatory system. Self-regulation involves such deliberate strategies as the operation of plans, arranging self-reward and self-punishment, and the calculation of the effort needed in order to achieve one's goals.

It is now appropriate to consider how psychotherapy affects these motivational processes and systems. Although the overall goal of therapy may be that of symptom reduction, achieving it is likely to involve the subgoals of raising expectancies of competence and exploring or eliminating potential goal conflict. Initially, the therapist will try to influence the patient's self-regulatory system to increase effort and persistence at jointly agreed tasks, such as planned exposure to feared situations. As has been argued, expectancies are likely to be raised by techniques such as the drawing up of graded hierarchies, with emphasis on the mastery of each step before progressing to the next one. Therapists also typically tell their patients that the methods have worked well for other people, and they might ask patients to recall other occasions on which they succeeded in overcoming fear. Effort may also be enhanced by emphasis on the benefits of eliminating the symptom or behavior, by social approval from a group of patients having similar difficulties, or by arranging contingent rewards for achieving targets.

From time to time, however, it will be difficult to obtain a satisfactory level of effort output, even though the patient reports wanting to get rid of the symptom or behavior and appears to agree on the method to be employed. Homework or therapy sessions may be missed, and attempts to strengthen self-regulatory behaviors fail. At times these counter-therapeutic behaviors will themselves be intentional and occur for reasons of which the patient is consciously aware. For instance, there may be a major difficulty that the patient perceives as the real problem and that he or she is unwilling to disclose until he or she feels more confident in the therapist. Alternatively, patients may be attending therapy because their doctor has told them their problem is psychological, even though they themselves see the origin of their symptoms as exclusively physical. At other times countertherapeutic behaviors will not be intentional but will be subjectively experienced by the patient as arising from lack of energy, forgetfulness, or the pressure of other demands on their time. These signs may indicate the influence of nonconscious cognitions representing past experiences of reward and punishment.

Obviously it will be of little value to ask the patient to introspect about the causes of this behavior. Instead it may be helpful to try and infer the nature of the cognitions from observations of behavior or explorations of the associative links between the nonperformed task and other known sources of punishment. It will only be possible to give two brief examples of these approaches here. Depressive behaviors such as sadness and tearfulness are often extremely effective in eliciting sympathetic reactions from others, and may be deployed in situations where the support or cooperation of others is uncertain. Such contingent relations are often noticeable to friends and relatives but not to patients themselves. Patients who have a history of being rewarded in this way for displaying negative emotions may have formed an

internal representation that these behaviors have a high probability of being followed by social rewards. This expectation of reward will exert an opposite influence on behavior from that of consciously held beliefs about the undesirability of experiencing these emotions.

This example makes use of a functional analysis of behavior and reinforcements familiar to behavior therapists. The second example makes use of a semantic analysis that will be more familiar to adherents of Kelly's personal construct theory. Many physical and psychological symptoms appear to be related to high levels of arousal, and such patients are usually encouraged to program more periods of relaxation into their daily schedules. A proportion of patients, however, although they agree that they ought to relax more, find it almost impossible to reduce the demands they perceive to be placed on them. When asked to provide associations to "relaxation" and related words, for instance by employing repertory grid methods, they may generate associations such as "lazy," "slapdash," or "careless," indicating an internal representation that relaxation is associated with negative rather than positive qualities. Once again this internal representation of negative consequences may exert an opposite influence on behavior from the consciously held intention to become more relaxed.

Rather than be satisfied with characterizing patients as unmotivated, behavior therapists should, therefore, extend their analyses of countertherapeutic behavior to encompass conscious and nonconscious expectancies, values, and goals. A major difference between this and the psychoanalytic approach, which also acknowledges nonconscious influences, is that in the latter resistance or noncooperation is seen as universal and inevitable, whereas in the former it is only expected and observed in a proportion of patients. This is likely to be due to the different aims and methods of the two forms of therapy. Another difference is that in psychoanalytic theory feelings are often considered to remain inaccessible to consciousness because they have been repressed, an assumption not generally made by cognitive behaviorists. In either approach, however, the presence of nonconscious influences is only inferred after careful analyses of behavior, and should lead to empirical testing within the therapy. It goes without saying that there are many other causes of lack of motivation that have more to do with the therapist or with the patient's social and interpersonal environment. Nevertheless, it is contended that a full theoretical account of patient motivation requires a consideration of intentional behavior and of behavior that is mediated by cognitions to which patients do not have conscious access.

ACKNOWLEDGMENTS

I am grateful to Dr. M. Power for helpful discussion of an earlier draft of this chapter.

REFERENCES

- Alloy, L. B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review*, *91*, 112-149.

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215.
- Bandura, A. (1978). The self-system in reciprocal determinism. *American Psychologist*, *33*, 344–358.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, *37*, 122–147.
- Bandura, A., Adams, N. E., Hardy, A. B., & Howells, G. N. (1980). Tests of the generality of self-efficacy theory. *Cognitive Therapy and Research*, *4*, 39–66.
- Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology*, *41*, 586–598.
- Bandura, A., & Simon, K. M. (1977). The role of proximal intentions in self-regulation of refractory behavior. *Cognitive Therapy and Research*, *1*, 177–193.
- Bargh, J. A. (1984). Automatic and conscious processing of social information. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of Social Cognition* (Vol. 3) (pp. 1–43). Hillsdale, NJ: Erlbaum.
- Barrios, B. A. (1983). The role of cognitive mediators in heterosocial anxiety: A test of self-efficacy theory. *Cognitive Therapy and Research*, *7*, 543–554.
- Beck, A. T. (1983). Cognitive therapy of depression: New perspectives. In P. J. Clayton & J. E. Barrett (Eds.), *Treatment of depression: Old controversies and new approaches* (pp. 265–290). New York: Raven Press.
- Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive Therapy of Depression*. New York: Wiley.
- Blatt, S. J., Quinlan, D. M., Chevron, E. S., McDonald, C., & Zuroff, D. (1982). Dependency and self-criticism: Psychological dimensions of depression. *Journal of Consulting and Clinical Psychology*, *50*, 113–124.
- Bradley, G. W. (1978). Self-serving biases in the attribution process: A re-examination of the fact or fiction question. *Journal of Personality and Social Psychology*, *36*, 56–71.
- Brewin, C. R. (1984). Attributions for industrial accidents: Their relationship to rehabilitation outcome. *Journal of Social and Clinical Psychology*, *2*, 156–164.
- Brewin, C. R. (1985). Depression and causal attributions: What is their relation? *Psychological Bulletin*, *98*, 297–309.
- Brewin, C. R., & Antaki, C. (1982). The role of attributions in psychological treatment. In C. Antaki & C. R. Brewin (Eds.), *Attributions and psychological change*. (pp. 23–44). London: Academic Press.
- Brewin, C. R., & Antaki, C. (1987). An analysis of ordinary explanations in clinical attribution research. *Journal of Social and Clinical Psychology*, *5*, 79–98.
- Brewin, C. R., & Shapiro, D. A. (1984). Beyond locus of control: Attributions of responsibility for positive and negative outcomes. *British Journal of Psychology*, *75*, 43–49.
- Conditte, M. M., & Lichtenstein, E. (1981). Self-efficacy and relapse in smoking cessation programs. *Journal of Consulting and Clinical Psychology*, *49*, 648–658.
- DiClemente, C. C. (1981). Self-efficacy and smoking cessation maintenance: A preliminary report. *Cognitive Therapy and Research*, *5*, 175–187.
- DiClemente, C. C., Prochaska, J. O., & Gibertini, M. (1985). Self-efficacy and the stages of self-change of smoking. *Cognitive Therapy and Research*, *9*, 181–200.
- Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology*, *29*, 288–298.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*, 215–251.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117–140.
- Frankel, A., & Snyder, M. L. (1978). Poor performance following unsolvable problems: Learned helplessness or egotism? *Journal of Personality and Social Psychology*, *36*, 1415–1423.
- Hammen, C., Marks, T., Mayol, A., & DeMayo, R. (1985). Depressive self-schemas, life stress, and vulnerability to depression. *Journal of Abnormal Psychology*, *94*, 308–319.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Johnson, D. A. W. (1981). Depression: Treatment compliance in general practice. *Acta Psychiatrica Scandinavica (suppl.)*, *63*, (Suppl. 290) 447–453.
- Jones, E. E., & Berglas, S. (1978). Control of attributions about the self through self-handicapping strategies: The appeal of alcohol and the role of underachievement. *Personality and Social Psychology Bulletin*, *4*, 200–206.
- Kanfer, F. H. (1970). Self-regulation: Research, issues, and speculations. In C. Neuringer & J. L. Michael (Eds.), *Behavior modification in clinical psychology*. New York: Appleton-Century-Crofts.

- Kassin, S. M., & Baron, R. M. (1985). Basic determinants of attribution and social perception. In J. H. Harvey & G. Weary (Eds.), *Attribution: Basic issues and applications* (pp. 37–64). Orlando, FL: Academic Press.
- Kazdin, A. E. (1979). Imagery elaboration and self-efficacy in the covert modeling treatment of unassertive behaviour. *Journal of Consulting and Clinical Psychology, 47*, 725–733.
- Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation. Vol. 15* (pp. 192–240). Lincoln, NE: University of Nebraska Press.
- Lane, T. W., & Borkovec, T. D. (1984). The influence of therapeutic expectancy/demand on self-efficacy ratings. *Cognitive Therapy and Research, 8*, 95–106.
- Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* New York: Appleton-Century-Crofts.
- Ley, P. (1977). Psychological studies of doctor–patient communication. In S. Rachman (Ed.), *Contributions to medical psychology* (Vol. 1, pp. 9–42). Oxford: Pergamon Press.
- Lick, J., & Bootzin, R. (1975). Expectancy factors in the treatment of fear: Methodological and theoretical issues. *Psychological Bulletin, 82*, 917–931.
- Mandler, G. (1984). *Mind and Body: The psychology of emotion and stress*. New York: W. W. Norton.
- Meier, S., McCarthy, P. R., & Schmeck, R. R. (1984). Validity of self-efficacy as a predictor of writing performance. *Cognitive Therapy and Research, 8*, 107–120.
- Morris, P. E. (1981). The cognitive psychology of self-reports. In C. Antaki (Ed.), *The psychology of ordinary explanations of social behaviour* (pp. 183–203). London: Academic Press.
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231–259.
- Posner, M. I., & Snyder, C. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium*. Hillsdale, NJ: Erlbaum.
- Pyszczynski, T., & Greenberg, J. (1983). Determinants of reduction in intended effort as a strategy for coping with anticipated failure. *Journal of Research in Personality, 17*, 412–422.
- Rhodewalt, F., Saltzman, A. T., & Wittmer, J. (in press). Self-handicapping among competitive athletes: The role of practice in self-esteem protection. *Basic and Applied Social Psychology*.
- Rotter, J. B. (1954). *Social learning and clinical psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*, 1–66.
- Shapiro, A. K. (1971). Placebo effects in medicine, psychotherapy, and psychoanalysis. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (pp. 439–473). New York: Wiley.
- Smith, T. W., Snyder, C. R., & Perkins, S. C. (1983). The self-serving function of hypochondriacal complaints: Physical symptoms as self-handicapping strategies. *Journal of Personality and Social Psychology, 44*, 787–797.
- Snyder, C. R., & Smith, T. W. (1982). Symptoms as self-handicapping strategies: The virtues of old wine in a new bottle. In G. Weary & H. Mirels (Eds.), *Integrations of social and clinical psychology* (pp. 104–127). New York: Oxford University Press.
- Strickland, B. R. (1978). Internal-external expectancies and health-related behaviors. *Journal of Consulting and Clinical Psychology, 46*, 1192–1211.
- Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.
- Wachtel, P. (Ed.). (1982). *Resistance: Psychodynamic and behavioral approaches*. New York: Plenum Press.
- Wilson, T. D. (1985). Strangers to ourselves: The origins and accuracy of beliefs about one's own mental states. In J. H. Harvey & G. Weary (Eds.), *Attribution: Basic issues and applications* (9–36). Orlando, FL: Academic Press.
- Wilson, T. D., Dunn, D. S., Bybee, J. A., Hyman, D. B., & Rotondo, J. A. (1984). Effects of analysing reasons on attitude–behavior consistency. *Journal of Personality and Social Psychology, 47*, 5–16.
- Wilson, T. D., Lassiter, G. D., & Stone, J. I. (1984). *Regulated versus unregulated nonverbal behaviour in social interaction: Evidence for limited access to mental states*. Unpublished manuscript, University of Virginia.
- Wylie, R. C. (1979). *The self-concept* (Vols. 1 & 2, rev. ed.). Lincoln, NE: University of Nebraska Press.
- Zuckerman, M., Larrance, D. T., Spiegel, N.H., & Klorman, R. (1981). Controlling nonverbal cues: Facial expressions and tone of voice. *Journal of Experimental Social Psychology, 17*, 506–524.

CHAPTER 14

Emotion, Cognition, and Action

Leslie S. Greenberg and Jeremy Safran

INTRODUCTION

In this chapter a theoretical framework that integrates ideas on affect, cognition, and action will be presented as a guide to thinking about the process of change in psychotherapy. We will briefly review some of the newer concepts in the psychology literature that inform our thinking about integration, present an integrative model of emotional processing, and discuss its clinical implications. The objective is not to articulate fully a comprehensive model of the relationship between the three domains of affect, cognition, and behavior. We aim, rather, to present a number of considerations relevant to the construction of an integrated model of functioning, in the hope that this will be an impetus to the development of further theory and much needed research relating these three different subsystems.

Central to our theoretical argument is the idea that the experience of emotion is the end product of a set of parallel, automatic, or unconscious information processing activities that involve the synthesis of subsidiary expressive motor, schematic, and conceptual level information (Greenberg & Safran, 1984a, 1986; Leventhal, 1979, 1982). Recent network analyses of emotion (Bower, 1981; Bower & Cohen, 1982; Bower & Mayer, 1986; Lang, 1983, 1984; Leventhal, 1982, 1984) that view emotions as being linked to expressive-motor reactions, autonomic reactions, memories, images, and ideas, suggest that eliciting any one of these components can evoke other parts of the network. From this perspective, the processes involved in emotion, cognition, and behavior, rather than being seen as separate and independent, are viewed as different components of the same structure joined together in a network. Activation of one of the components automatically spreads to the other components of the network, increasing the probability of particular responses and emotional experience. This automatic evocation of previously unavailable parts of the network has important implications for the practice of therapy. A central goal of intervention becomes

Portions of this chapter are adapted from Safran and Greenberg (1986).

LESLIE S. GREENBERG • Department of Psychology, York University, North York, Ontario M3J 1P3, Canada. JEREMY SAFRAN • Clarke Institute of Psychiatry, Toronto, Ontario M5T 1R8, Canada.

one of evoking those organized associative structures which govern emotional experience. Evocation of these structures can lead directly to behavior change, and in addition can make the associative structures themselves amenable to restructuring.

Until recently, affect, cognition, and behavior have generally been viewed as separate and independent processes that are causally related in a linear fashion rather than as interdependent processes operating together. In addition, the emphasis in any particular linear view has generally been on the relationship between only two of the three components. For example, it has been argued on the one hand that change in attitude or cognition leads to change in behavior, and on the other that change in behavior leads to change in cognitions or emotions. Similarly, it has been argued that change in cognition leads to change in emotion and conversely that emotional change leads to cognitive change. All viewpoints have received some empirical support (Greenberg & Safran, 1987).

Experimental investigation of psychotherapy and behavior change has shown that performance based treatments and more recently cognitive modification treatments bring about both cognitive and behavioral change. The role of the evocation and processing of emotion in producing therapeutic change has not been extensively researched. It appears, however, that emotion can be used in a variety of ways to produce therapeutic change (Greenberg & Safran, 1984a, 1987). Clinical observation suggests that unwanted emotions, such as grief and fear, can be reduced by fully experiencing and reprocessing the emotion (Perls, Hefferline, & Goodman, 1951; Rachman, 1980), whereas other emotions, such as anger and sadness, can be therapeutically evoked in order to motivate behavior change and adaptive problem solving (Greenberg & Johnson, 1986a, b; Greenberg & Safran, 1986). It appears therefore that the development of a comprehensive approach to psychotherapy and behavior change would be enhanced by a model of functioning that takes into account behavioral, cognitive, and affective processes.

A number of integrative concepts have recently been put forward in the psychological literature to account for emotional experience and the cognitive organization of complex stimulus and response patterns related to emotion. The information-processing literature has proposed associative and semantic network theories of emotion in an attempt to conceptualize the interplay between emotions, cognitions, and behavior (Bower, 1981; Bower & Cohen, 1982; Lang, 1983, 1984). Schema theory continues to develop in cognitive psychology in an attempt to account for cognitive organization and emotional processing (Leventhal, 1984). In addition, evaluative conditioning theory (Levey & Martin, 1983; Martin & Levey, 1978) has recently been proposed in the behavioral literature with the aim of extending a conditioning model to account for the subjective evaluation of stimuli. The characteristics of these different integrative approaches and the manner in which they combine affective, cognitive, and behavioral processes will be briefly reviewed in the following sections.

ASSOCIATIVE NETWORK MODELS

As we have argued elsewhere (Safran & Greenberg, 1986), semantic network theories are essentially an evolution from a basic associationist model. The basic assumption in these models is that words and concepts vary with respect to the degree

of associative relationships they have with one another. Semantic network theories differ from simple associative theories by going beyond a simple model of word association to a concept of a complex network of associations that provide meaning to the elements of the network.

In semantic network models, the relationship between different items stored in memory is represented in terms of the degree of association between items in the network (represented by the length of the link or pathway), the nature of the association (represented by the label of the link), and the nature of the items that are associated (represented by the labels attached to nodes in the network). Different network theories share a number of assumptions and are different with respect to certain assumptions. A common feature of all semantic network models, however, is that they provide a model of the organization of information in memory that takes into account that there are different kinds of information stored in memory and different kinds of associations between this stored information. The notion of spreading activation is central to all semantic network theories and to all attempts to make predictions about different characteristics of network functioning, such as the speed with which the activation of one element in a semantic network should activate other elements in the network.

Lang (1979, 1983, 1984) has proposed an information-processing theory of emotion that maintains that affective dispositions to act are coded in memory in propositional form. These propositions are organized into networks related by association, that is, emotions are represented in memory as networks of information, in which response, stimulus, and meaning information is associatively connected. According to Lang (1983), emotion is best regarded as "an action set, defined by a specific information structure in memory, which when accessed is processed as both a conceptual and a motor program" (p. 7). These information structures, referred to by Lang as emotion prototypes, contain three types of information: (a) information about stimuli or events that have elicited the relevant emotion in the past; (b) information about the complete emotional response itself (including verbal behavior, overt acts, and relevant visceral and somatic changes); and (c) information defining the meaning of the eliciting stimulus and the emotional response.

An emotion prototype becomes automatically activated and processed when an individual attends to information that matches a number of the coded propositions. For example, the individual who is exposed to environmental stimuli that are sufficiently similar to those stimuli that previously evoked a particular emotion in the past may automatically undergo an activation of the relevant emotion prototype. The probability that a particular emotion prototype will be activated increases as the number of coded prototype features that are matched increases. Lang theorizes, however, that some prototype features or propositions may be more central or critical in this process than others. Thus one particular image may be more critical in this accessing process than another, or a particular motor response may be more critical than another. The important thing to bear in mind, however, is that once the right proposition or combination of propositions are matched, the entire emotion prototype is automatically activated and processed, and the individual automatically experiences an emotion consisting of autonomic, expressive-motor, imagery, and conceptual components.

In Lang's theoretical perspective, activation of the prototype provides the organism with information that motivates action. As Lang (1983) points out:

Activation of an emotion prototype always prompts efferent outflow whether or not overt behaviour is occasioned. The efferent activity is the output of motor sub-routines that are linked to the deep structure of response information. The emotion itself is an action set. . . . Emotions are always about doing something, and setting, meaning, and pattern of action are all coded in the same associative network. (p. 32)

Lang thus proposes that emotions are to be understood as behavioral acts and stresses that emotions are cognitively represented as response information. Because affect and cognition contain a significant motor component, a focus on the motor system appears to provide interesting possibilities for the investigation of the interaction of affect and cognition.

An important aspect of the network model is the manner in which emotions are cognitively represented as response information. Affect, cognition, and behavior are no longer viewed as separate and independent. They are seen to operate in an integrated and interdependent manner with response information and the motor system being an integral part of the network, and emotion being the experience of the activation of the network. Internal cognitive representations of prior expressive-motor responses are linked with features of the stimulus situation to constitute internal cognitive structures. Emotion is brought about by the activation of those structures that contain expressive motor responses as an efferent component of the network. Appropriate stimulus presentation activates the network releasing the response aspects of the network, thereby producing the experience of emotion.

Recently, Bower and colleagues (Bower, 1980; Bower & Cohen, 1982; Bower & Mayer, 1986) have been utilizing network models to explain the relationship between emotion and retrieval and emotion and learning. Bower's initial findings (1981) suggested a strong mood-dependent retrieval effect in which retrieval increased with the degree of similarity between a subject's mood at learning and at recall. Association in the network between learned material and mood was assumed to be formed by contiguity. In addition to a mood-dependent retrieval effect, Bower also reported a mood-congruity learning effect. As a demonstration of the later effect, subjects were shown to learn items that were mood congruent more efficiently than items that were mood discrepant. Later studies on the effect of mood on memory have been unable to replicate the earlier findings on the retrieval effect (Bower & Mayer, 1986). Although the learning effect was replicated, these later studies of the retrieval effect showed that the effect occurred only when subject's causally related their aroused emotional responses to the learned materials. Thus the earlier assumption of learning by contiguity has been replaced by a causal belongingness hypothesis in which a strong association is formed in memory between a stimulus event and the emotion it evokes only if subjects causally attribute their emotional reaction to the stimulus. This hypothesis seems far more plausible than the simple contiguity hypothesis, as it takes the subjects encoding of the situation into account.

In these network models emotions are thus seen to serve an important organizational role in memory. Different models have been proposed in which different cognitive, affective, and behavioral components are seen as being associated in the network as nodes. Emotion can be evoked through the activation of any of the units in memory that are linked to the emotion node. For example, the activation of a specific episodic memory associated with a specific emotion can activate that emotion. Similarly, the activation of a particular expressive behavior linked to a specific emotion

node can activate that emotion. The activation of any one particular unit in memory that is linked to a specific emotion node may not necessarily activate that specific emotion node. It may, however, raise the threshold of that particular node, so that it will be more easily activated by another unit in memory that is linked to that emotion node.

SCHEMA MODELS

The schema concept proposed originally by Bartlett (1932) and used extensively in developmental psychology by Piagetians, has now become prominent in the cognitive psychology literature. Although specific definitions of the schema construct still vary from theorist to theorist, there is general agreement on the principle that cognitive structures in memory (schemata) organize information abstracted from prior experience and guide the processing of new information and the retrieval of stored information. As Fiske and Linville (1980) pointed out, the schema construct, when evaluated against specific criteria of good scientific theory, shows definite theoretical merit. The schema construct has been demonstrated to have some predictive utility, has been shown to be profitable in terms of generating productive research, and appears to generalize to a large number of domains of interest in psychology (e.g., memory, attention, social psychology, and clinical psychology).

Leventhal (1982) recently employed the schema construct as a way of understanding the relationship between emotion and other cognitive processes. He hypothesized that emotional experiences are schematic structures consisting of at least three levels of information processing (expressive-motor, schematic memory, and conceptual) that combine to create the experience of emotion. These schemata include (a) expressive-motor behaviors and autonomic patterns associated with specific emotions, (b) episodic memories and images of specific events associated with a specific emotional experience, and (c) conceptual rules and propositions about how to respond to specific situations and the possible consequences of these responses. Leventhal (1982) proposed that these schematic structures are coded in memory at a preattentive level. In his perceptual-motor processing model of emotion the activation of one component of a specific emotion schema increases the potential for activating the entire schematic structure. Subjective emotional experience is the effect of the activation of a schematic structure.

These schematic structures are quite similar to Lang's emotion networks or prototypes and consists of similar cognitive, affective, and behavioral constituent elements. Leventhal and Lang also agree that the activation of one element of an emotion prototype, or the right combination of critical elements, will activate the entire prototype and result in the subjective experience of the associated emotion.

THE EVALUATIVE CONDITIONING MODEL

Martin and Levey (1978) have suggested that the establishment and transfer of an evaluative response to a previously neutral stimulus is the carrier of the mechanism of classical conditioning. In their approach they suggest that what is primarily

learned is a central evaluation state that is triggered by an evaluative response to salient stimuli in the presence of previously neutral stimuli. Arguing that the evaluation of positive and negative valences in the world is an imperative and innate feature of all forms of life, they posit that the evaluative response is the characteristic internal reaction of all organisms to environmental stimuli. This response, which is essentially an evaluation in terms of good/bad, liked/disliked, pleasant/unpleasant, is accompanied by a pattern of physiological change organized toward approach or withdrawal. The evaluative response is a subjective, unmediated, response elicited by salient stimuli and transferred to contiguous neutral stimuli.

Two types of learning are suggested by Levey and Martin (1983). One involves classical conditioning and the evaluative response: this is referred to as learning "rules of consequence." The other, learning of "rules of sequence," is a cognitive learning process in which regular and recurring sequences are coded and stored in accessible form.

In this view, conditioning occurs only if an evaluative response is first elicited and what is conditioned is the evaluative response itself. The consequent behaviors are determined by the situation in which the response is evoked and are essentially open ended. It is not the behaviors but the evaluative response that has been conditioned. The learning of rules of consequence allow organisms to learn regularities in the environment and to evaluate an object in the environment in terms of a prior negative experience with it. Thus an animal reacts with a negative evaluative response to an insect as if the insect were the unpleasant flavor it experienced as a consequence of previously biting this insect. It has negatively evaluated the flavor and this negative evaluation has been transferred to the insect. Martin and Levey's view that something like affective appraisal of environmental events must be present in all organisms is held by a number of theorists of diverse orientation, although differences exist in the degree of mediation posited by the various theorists. Lazarus (1984) for example, favors a more mediated view whereas Arnold (1960) and Zajonc (1984) favor unmediated views.

Levey and Martin's (1983) second type of learning involves rules of sequence that refer to the temporal processing of environmental sequences. This form of learning maps the course of events over extended periods of time and is concerned with the structure of the environment with long-term outcomes, and with sequences of events. Essentially, this is a cognitive summary of experience often using language to code information about the temporal order and to store such information for subsequent use.

In summary, the authors propose that adaptive behavior is largely concerned with the apprehension of sequences of events and learning the rules of these sequences. The rules fall into two categories: those concerned with isolated events and their immediate consequences and those concerned with larger sequences and their outcomes.

Considering applications of their view, Levey and Martin (1983) suggest that there are three levels at which maladaptive behavior can be approached: (a) the level of affective evaluation; (b) the level of cognitive structure in which the rules of sequence are summarized (these are concerned with a "knowing" component as opposed to doing); and (c) the level of planned strategies that involves the action component and is based on knowing the consequences of sequences of actions. They

suggest that treatment methods and combinations should be based on an assessment of which level is primarily involved in the maladaptive behavior.

COMPARISON OF MODELS

It appears that a number of the constructs in the previous three models are quite similar. Specifically, they all suggest that conscious evaluation of stimuli is not necessary for the stimuli to have effects and that different events or features that are associated either temporally, semantically, or conceptually are linked in memory in one fashion or another. Thus according to all these models, information can be coded and linked in memory at a preattentive level, out of awareness. They all stipulate that the activation of one unit in memory can lead to the activation of associated units in memory, and that this activation can all take place at the preattentive or unconscious level. Thus, in accordance with Zajonc's position on the primacy of affect, it appears that conscious appraisal of stimuli is not necessary for an emotional response. Finally, they all suggest the importance of three levels of processing: immediate appraisals involving expressive-motor and physiological responses, schematic and semantic processing in which experience is combined and stored, and conceptual and propositional processing in which conceptual rules are formed.

Specific differences in terms of the predictions that can be derived from schema theory, associative network theory, and evaluative conditioning theory will need to be spelled out. There is, however, sufficient commonality among them at an explanatory level to suggest that they form a basis for a new integrative view of the relationship between emotion, cognition, and behavior.

AN INTEGRATIVE MODEL OF EMOTIONAL PROCESSING

A number of the major tenets of the integrative model of emotional processing that we have recently proposed (Greenberg & Safran, 1984a, b, 1987; Safran & Greenberg, 1986), are outlined in the following sections. This approach draws on and is consistent with principles outlined earlier in the network, schema, and evaluative conditioning models.

THE EMOTIONAL SYNTHESIS PROCESS

We suggest that the basic structure for emotional experience appears to be provided by a central neural program, which is wired-in rather than learned. The central neural program can be thought of as a template for emotional experience. In response to both external and internal events, it generates neural impulses that activate expressive-motor behaviors and lead to the subjective experience of emotion.

There is some disagreement among theorists as to which emotions are primary and which emotions are more complex derivations of these primary emotions. There

is agreement, however, amongst all theorists assuming a biological/evolutionary perspective on emotion, that the structure for certain primary core emotions is wired into the human organism. Consistent with theorists such as Arnold (1960) and Leventhal (1979, 1982) we hypothesize that the wired-in neurological substrate for emotional experience includes a code for specific configurations of expressive-motor behaviors that correspond to specific primary emotions—including at least the six emotions with identifiable facial expressions: fear, anger, sadness, surprise, disgust, and joy (Ekman & Friesen, 1975). We are in no sense, however, claiming that emotional experience in the adult human being is in any sense restricted to these simple, primary emotions and associated expressive-motor configurations. Rather, the basic neurological template for emotional experience becomes elaborated in the human being into subtle blends of emotional experience such as pride, envy, and humility that are characteristic of human functioning.

Based on Leventhal (1979, 1982) we hypothesize that complex human emotional experience comes about by a type of automatic information-processing activity that synthesizes information generated through perception of the environment, neural impulses associated with expressive-motor behaviors, schematic memories, and higher-level conceptual processing. Emotion is thus constructed or synthesized by a set of complex information processing activities.

In this model, the human organism is seen as responding to the environment in an immediate, reflexive fashion and the type of immediate appraisals or subjective evaluations of the environment that are made relate to biological and psychological survival. In our view the evaluative response of good/bad is an important one but is not the only primary appraisal of environments. As well as the evaluative response we suggest the existence of a number of other appraisals relating to organism-environment interventions (see for example Plutchik, 1980). Appraisals relating to attachment and separation (Bowlby, 1969) and novelty (Berlyne, 1960; Butler, 1965) are probably as fundamental as the good/bad response. Histories of negative experiences in the domains of these different appraisals probably relate to different dysfunctions. The important point is that although evaluative conditioning is consistent with the model we have proposed, we posit the existence of more than one type of primary appraisal. Although the nature of these other appraisals is more speculative, we believe that it is important not to lose sight of the possible existence of more than one type of primary appraisal in explaining subjective human experience.

We hypothesize, therefore, that people engage in immediate perceptual-motor appraisals of environmental events. Such appraisals activate networks, the outputs of which become synthesized into subjective emotional experience. This perceptual motor or primary appraisal is automatic and does not involve any conscious conceptual appraisal. At the same time, however, that the automatic appraisal is taking place, higher level conceptual processing is appraising the primary appraisal (Safran & Greenberg, 1982). This multilevel, parallel processing of information, in which information generated from inside and outside of the organism is constantly being appraised, is integrated out of awareness and results in the conscious experience of emotion. The information generated by this synthetic process produces ongoing emotional experience and becomes stored in memory in a tightly associated structure.

Thus, from birth, a child is accumulating memory stores consisting of episodic memories, images of eliciting environmental events, evoked expressive-motor responses, associated autonomic arousal, and associated conceptual appraisal.

These memory structures can be viewed either as schemata (Leventhal, 1979, 1982) or as networks (Lang, 1983). Emotional experience thus becomes coded in memory structures that incorporate a number of subsidiary components. When an individual either attends to information or generates information internally that matches one of the subsidiary components, the probability of other associated components becoming activated increases. The information generated by the activation of these cognitive-affective behavioral networks results in the conscious experience of emotion.

In our view emotional experience is not, however, restricted to the primary biologically based categories of emotion; rather, complex and subtle derivatives of the more basic emotions substructures are established through the development of the complex cognitive-affective-behavioral structures in memory that store the individual's unique experiences in life, and his or her responses to them.

ADAPTIVE ACTION TENDENCIES

In our view, the physiological and expressive-motor behaviors associated with emotion are best thought of as action tendencies in the world. These action tendencies or dispositions may or may not be transferred into goal directed behavior depending on decisions executed by higher level processing.

From an evolutionary perspective, the type of emotion system that we have described has the advantage of rapidly activating adaptive behavior in response to an event that requires action. Thus, there is an immediate, rapid perceptual appraisal of situations that gives rise to physiological and expressive-motor responses that are action tendencies in the world. Human beings are adaptive organisms constantly adjusting themselves to and attempting to master the environment. Incoming stimuli are appraised and automatically activate action tendencies toward the world. Thus an individual is continually preparing to move toward and take in, or to eliminate, or to attack, or to move away from some aspect of its environment.

A number of emotion theorists (Arnold, 1960; Izard, 1977; Leventhal, 1982; Plutchik, 1980; Tomkins, 1980) argue that emotional processing has evolved in the human species through a process of natural selection and that emotional processes play an adaptive role in human functioning. This is not to say that all emotion is necessarily always adaptive in all contexts, but rather that emotional processes do play an adaptive role in human functioning. Emotions appear to provide us with information about ourselves as organisms in interaction with the environment. They thus constitute a bridge through which people are linked to their ecological niche (Gibson, 1979). This bridge has evolved through a process of natural selection. Thus emotions are not epiphenomena or the product of the cognitive interpretation of arousal, as Schachter and Singer (1962) theorize. Rather, the basic structure for specific emotions is hardwired into the human animal (Izard, 1977; Leventhal, 1982; Tomkins, 1980).

According to Lang (1984), emotions function as action dispositions. Different classes of action are inherent in different emotions. Anger, for example, will lead to aggressive, self-protective behavior if carried through into action. Fear will lead to self-protective behavior through flight, if carried forward into action. Loneliness can lead to affiliative action. Affiliative behaviors play a strong role in terms of the survival of the species. Love can lead to affiliation and to procreation.

In a situation in which an individual is physically threatened, a neural impulse that organizes the person for fight or flight is generated by the central neuromotor template. Certain expressive motor behaviors might be generated, such as muscle tension in the legs preparing the individual to run or a fighting posture preparing the individual to defend her or himself. A particular pattern of muscular patterning in the face might be simultaneously generated. This facial patterning is another type of expressive-motor behavior action tendency in the world that also serves a communication function.

The example described earlier is the most simple danger situation and is generally more relevant for infrahuman species than it is for human beings. For people, however, complex variations or elaborations of this prototype of safety and danger arise. For example, a person who experiences a blow to his or her self-esteem may increase their level of arousal to refute the criticism and may generate facial-muscular patterning associated with anger, signaling to the offending individual to be more cautious. Thus with human beings, the type of situations to which these expressive motor responses are relevant is typically less related to immediate life and death situations and more related to complex social interactional themes. Our position is that autonomic arousal, various physiological changes, specific facial expressions, and various patterns of muscular tension in the body, are all part of the action tendency that is generated.

SPECIFIC PATTERNING AND SOMATIC FEEDBACK

As previously stated, the schematic coding for an emotion includes the eliciting event, the subjective experience, associated images, and the accompanying expressive-motor behaviors and autonomic reactions. Thus, although there is little empirical evidence indicating that specific patterns of autonomic activity are reliably associated with specific emotions across individuals (Greenberg & Safran, 1987), it is entirely possible that for a given individual complex patterns and configurations of expressive-motor and autonomic reactions become coded in schematic memory over time. In this way they become consistently associated with specific emotional states for that individual. For example, a particular individual may consistently respond to situations of threat with an idiosyncratic response pattern experienced as a hollowness in the stomach, shaky legs, and cold hands, whereas anger may be experienced as feeling hot and prickly with an increased heart beat. Another individual may experience different idiosyncratic responses to threat and anger, unique to him or her.

According to Leventhal (1984), Bower (1981), and Lang (1984), the elements of the emotion schema are coded together in a common memory locus. They are thus tightly linked together, and the activation of one component of an emotion

schema will activate the other components of that same schema. For this reason attending to the particular combination of expressive-motor behavior and autonomic arousal typically associated with a specific emotion for that person can elicit other associated components of that emotion. Thus clients can be instructed to pay attention to their internal sensations and to their expressive behaviors in order to evoke emotional experience. Becoming aware of a dryness in the throat, a clenched fist, or a tapping foot all help toward activation of the relevant affective schema.

Although the experience of emotion is not necessarily dependent on somatic feedback (Greenberg & Safran, 1987), somatic feedback can and often does contribute to the experience of emotion. Although the central neural template generates a neural impulse that feeds directly into subjective awareness, the information provided by this neural message is typically supplemented by somatic feedback. The fact that somatic feedback typically does contribute to the subjective experience of emotion can have important therapeutic implications.

Although somatic feedback may not be a necessary subsidiary component in the synthesis of emotions, attention to relevant somatic feedback may provide one starting point when attempting to synthesize adaptive emotions that are typically not synthesized because of maladaptive social learning. For example, an individual who is not synthesizing the emotion of sadness may begin to synthesize this emotion of sadness by attending to the associated feeling of heaviness in his or her facial muscles. Similarly, an individual who is not synthesizing the emotion of anger may begin to synthesize this emotion by attending to clenched muscles in the jaw, and a clenched fist.

HIGHER LEVEL PROCESSING

In addition to generating rapid and immediate action tendencies, the emotional synthesis process allows for the mediation of these action tendencies through more sophisticated information-processing activity. This higher level information processing elaborates action tendencies into emotions and these emotions in turn motivate rather than cause adaptive behavior.

Emotions are elaborate cognitive-affective complexes linked to action. Subjectively experienced emotions are the conscious awareness of the synthesis of different types of information that have been combined to create the opportunity for generating complex and flexible action plans. Emotions provide human beings with biological feedback that allows them to adapt to their environment through set goal patterns (Bowly, 1969). Human beings have a flexible response system rather than the fixed action pattern system found in many infrahuman species. The behaviors associated with emotional responses are thus not fully fixed but are mediated by more complex processing activity and different behaviors will result depending on how the internal response is processed. Although the wired-in nature of the emotional system allows the opportunity for immediate, adaptive, reflexive responses to environmental events, there is not an inflexible link between environment and behavior. The conceptual aspect of the emotional synthesis process creates a break in the environment-behavior chain. The emotional synthesis process thus generates action disposition information

that is subjected to further processing that ultimately may lead to action. Dysfunction can thus occur at two levels—the type of action disposition information generated or the type of processing this information undergoes.

EMOTION IN PSYCHOTHERAPY AND BEHAVIOR CHANGE

As we have argued previously (Greenberg & Safran, 1984a, b, 1987; Safran & Greenberg, 1986) a theoretical understanding of the relationship between emotion, cognition, and action such as the one suggested here can be extremely useful for purposes of understanding the development of emotional dysfunction as well as clarifying our understanding of the process of change in psychotherapy. An understanding of the function of emotion and the way in which the cognitive-affective behavioral system functions under optimal conditions provides us with clues as to the way in which this system can break down when people have emotional problems.

CLINICAL IMPLICATIONS

A central tenet of the perspective outlined here is that primary emotions generate important information about the meaning of events for us as biological organisms and motivate behavior in a potentially adaptive fashion. It is important to recognize that affect is information in the form of an action disposition. As Leventhal (1982) maintains, emotions play a role in providing us with information about the readiness of our biological machinery to interact with specific events in the environment, and in integrating abstract cortical functions with perceptual-motor reflexes to enable us to sense, think, act, and feel in an integrated fashion. Emotions

can be regarded as a form of meaning. They have significance for the person experiencing and expressing them. Their meaning has two aspects: they "say" something about our organismic state (i.e., they meter its moment-to-moment readiness), and they "say" something about the environment. (Leventhal, 1982, p. 122)

It follows that individuals who, for whatever reasons, are not able to use fully or do not have complete access to this information will function in a less than optimal fashion. A common clinical problem, in our observation, occurs when clients fail to synthesize fully certain adaptive emotional experiences, bearing in mind that emotional experience, as used here, incorporates cognition and action. Because of past experiences they may learn that it is inappropriate or dangerous to have certain types of emotional experiences, and as a result may restrict the expression of certain emotions or may even fail to synthesize completely certain types of emotions. We hypothesize that both intensity and degree of redundancy of specific classes of learning conditions play roles in determining to what extent an individual will have difficulty in synthesizing associated emotions. In more extreme maladaptive learning situations the individual may fail to develop in memory any elaborated representations of the relevant emotions and thus actually have difficulty in completely synthesizing the relevant emotional experience. When the maladaptive learning is less extreme, the individual may be able to synthesize partially the relevant emotion, but may have difficulty in fully experiencing and expressing it. Common areas of emotional deficit

are found in the inability to synthesize experiences of weakness or vulnerability, and experiences of anger (Greenberg & Safran, 1984a, 1987; Safran & Greenberg, 1982).

Notwithstanding the biologically adaptive function of emotion, there are certain instances in which emotional responses have become dysfunctional or maladaptive. There are a variety of times in therapy where one arrives at maladaptive affective responses that are highly entrenched and present major problems. For example, a type of conditioning process may have led to certain conditioned evaluations and associated anxiety responses. In this situation modification of the evaluation is required.

In our view, it is possible to modify evaluations and the associated expressive motor components of a network using a variety of procedures drawn from the practice of behavior modification. Procedures such as exposure or practice and reinforcement are still applicable in this view. The main difference is that a simple conditioned response or set of responses is not viewed as the source of the maladaptive emotional response. We regard the conditioning process as involving the conditioning of subjective evaluations, and the resulting expressive-motor responses as being associated through schematic processing with a set of other components in a total network. It is this network involving the subjective evaluation that is the source of the maladaptive reaction. Rather than a simple response conditioning process through associational learning, we envisage a more complex evaluative conditioning plus schematic learning process in which a number of components are bound together into a schema or associated in a network. We hypothesize that it is the primary appraisals and the link between components that requires change. Schemata that require modification are very primitive and often, because of their maladaptive nature, have been isolated from further experience and have not undergone much further elaboration or learning. Thus schemata that generate primary fear in response to objectively safe situations have faulty conditioned evaluations at their core. They can be modified by evoking the network and exposing it to new experience in order to modify the evaluation and some of the associative links in the network.

Some of the advantages of this perspective over a simple classical conditioning view of modification can be seen in the procedures suggested by the network view. First, this perspective suggests that activation of the network is required before any modification of it can take place. The importance of the evocation of the maladaptive affective experience is thus implied by this view in a way that it is not in a classical conditioning view. Second, it is not simply exposure or relaxation paired with an undesirable response that leads to modification but rather the admission of new information to the processing system that produces change by forcing a reorganization in the schemata or networks. This leads to an emphasis on different ways in which information is represented rather than assuming a simple relationship between stimulus and behavioral response.

INTEGRATED INTERVENTION

An integrated view of intervention flows naturally from an integrative theoretical model. Reciprocally interactive and mutually enhancing effects of behavior modification, cognitive modification, and affective modification interventions are suggested by our model. If, in fact, appraisals of stimuli, motor and behavioral responses,

sensory experience and verbal, conceptual representations are all linked in memory by complex networks, then change in one set of components will lead to change in the network. If these networks are activated through the priming of subsidiary components, it appears that the behavioral, cognitive, and affective responses of patients in therapy are much more interconnected than separate. Rather than a linear causal sequence in which change in behavior is viewed as leading to change either in cognition or change in affective experience or the more complex view that performance-based treatments are cognitively mediated, we suggest a truly reciprocal view of intervention. In this integrative view, intervention and change in any of the response systems can lead to change in any of the other response systems. Thus affective change, behavioral change, and cognitive change are circularly related and reciprocally determining.

It is important at this point to note a conceptual and linguistic difficulty. Affective, cognitive, and behavioral systems as are not ultimately independent entities in their own right. Rather all are linked together and fused. We have argued, along with others, that emotion is not independent of, or prior to, cognition or behavior but rather that emotion includes expressive-motor, schematic, and conceptual components. It is thus a blend of sensory, motor, and perceptual/conceptual elements. Clinically, however, many useful attempts have been made to describe and conceptualize a set of interventions and of patient problem states as operating at either behavioral, cognitive, or emotional levels. This possesses heuristic value and a certain face validity when dealing with clinical phenomena. Clinicians do track patient's behaviors, thoughts, or feelings separately and can distinguish them for practical purposes. Although they may not be as distinct at a theoretical level there does exist at the applied level a repertoire of interventions that have become known and can be described as either behavioral, cognitive, or affective. When we refer to the mutually interactive effects of the different response systems, it is the traditional, applied-level distinction that we are using.

Thus when we speak of modifying overt avoidance behaviors by means of exposure (i.e., a behavioral intervention), we are suggesting that links between networks nodes involving efficacy expectations as well as those between nodes involving primary appraisals, physiological and expressive-motor responses, etc., are modified. When emotions such as anger or sadness are evoked in therapy (i.e., an affective intervention) behavioral changes can be brought about by the activation of action tendencies, such as assertiveness or comfort seeking, whereas cognitive changes can be brought about by the influence of mood on self statements and memories. Changes in emotion thus lead to changes in cognition and new action dispositions. Similarly, changes in cognitions can lead to changes in behavior and changes in emotion. For example, when catastrophic expectations are realistically inspected, internal experience can be modified and behavior change promoted.

The clinical issue following from this perspective is probably best phrased in terms of which intervention in which domain is most appropriate at what particular time with what types of disorders (Greenberg, 1986; Rice & Greenberg, 1984). Clearly we are implying that the behavioral cognitive and affective systems are not independent (Greenberg & Safran, 1984, 1987). They do not require separate modification as suggested by some (Rachman, 1981), but rather certain types of intervention, by

virtue of the components of the network that they evoke, are best for evoking the network and making it accessible to modification.

It appears from reviewing the literature on psychotherapy and behavior change that the least clearly articulated forms of intervention are those in which emotion is utilized to bring about therapeutic change. We suggest the addition of a number of affective change methods to the existing repertoire of behavioral and cognitive methods.

EMOTIONAL CHANGE PROCESSES

The following categories of emotional change processes describe some of the therapeutic interventions that involve working with emotion:

1. Evocation of emotion to motivate new behavior
2. Emotional restructuring in which networks are evoked in order to change linkages in the network
3. Accurate acknowledgment of affective responses to provide response information and enhance problem solving

Each method involves different processes, has different goals, and produces different types of change. In the first category, evocation of emotion to motivate new behavior, emotion is aroused in the session as a means of changing behavior (Greenberg & Safran, 1987). Thus the experience of anger leads to the promotion of assertiveness (Kahn & Greenberg, 1980), sadness to the promotion of contact and comfort seeking behavior and the experience of certain fears to the promotion of less aggressive and more affiliative behaviors in marital interactions (Greenberg & Johnson, 1986a, b). In this emotional change process, the action tendency associated with the experience of emotion is elaborated into the behavioral response that the client is lacking, such as comfort seeking, or the response that would help alleviate the assessed distress, such as assertion. The action tendency along with the emotion is evoked by a variety of emotional stimulation methods, both verbal and nonverbal. Often nonverbal means of stimulation, such as use of imagery, enactments, music, and drawing are most effective, but verbal means, such as repetition and exaggeration of certain key phrases by the client, can also be effective (Greenberg & Safran, 1987).

In the second intervention category, emotional restructuring, the underlying response program needs to be accessed and run in order to make it amenable to restructuring. As Lang (1983) has pointed out, the more the stimulus configuration matches the prototype or internal structure, the more likely it is that the network will be evoked. Experiencing the emotion of fear or sadness or anger is necessary in order for this experience to be modified. *In vivo* exposure to feared stimuli is far more likely to evoke the fear response and the possibility of emotional restructuring than exposure to a picture. In treating phobias it is often more possible to create *in vivo* experiments than when dealing with other problematic conditions, such as depression, loss, or chronic anger problems. In this situation the therapeutic situation needs to be used as a laboratory for evoking and reprocessing reactions in order to restructure the cognitive-affective-behavioral network or schema. Interventions ranging from the use of imagery to evocative responding (Rice & Saperia, 1984) to enactments and

gestalt two-chair dialogues (Greenberg, 1984; Greenberg & Safran, 1987) help access and set in motion the cognitive and behavioral response patterns that need to be modified. Once the network is amenable, different types of interventions can be used to produce restructuring (Greenberg & Safran, 1987).

The principles underlying the final category, (i.e., acknowledgment of affect), are based on the adaptive nature of primary affective responses and on the importance of accessing this information to aid problem solving. Organisms that ignore their own affective feedback are not well situated to behave adaptively. Accurately acknowledging affective responses that were previously avoided or misinterpreted makes certain reactions and moods more understandable and provides new information to guide the individuals actions toward need gratification and goal attainment. Acting in the world to satisfy certain needs and wants is the sine qua non of competence and satisfaction. Without acknowledgment of feelings and desires, people lack the awareness of the action tendencies to motivate action.

CONCLUSION

Affect, cognition, and behavior are essentially fused. All interactions with the environment involve affect and cognition. There is no affective behavior without cognitive behavior. Rather, all behavior involves affect and cognition. An integrative theoretical model has been presented in which affect, cognition, and behavior are seen as interdependent aspects of human functioning. We suggest an integrative model of clinical intervention in which the clinician works at the applied level with what the patient is thinking, feeling, and doing and views attempts at cognitive, behavioral, and affective modification as reciprocally interactive and mutually enhancing.

REFERENCES

- Arnold, M. B. (1960). *Emotion and personality*. New York: Columbia University Press.
- Bartlett, F. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Berlyne, D. E. (1960) *Conflict, arousal and curiosity*. New York: McGraw Hill.
- Bower, G. H. (1981). Mood and memory. *American Psychologist*, 36, 129-148.
- Bower, G. H., & Cohen, P. R. (1982). Emotional influences on learning and cognition. In M. S. Clarke, & S. T. Fiske (Eds.), *Affect and cognition: The 17th annual Carnegie symposium on cognition* (pp. 263-289). Hillsdale, NJ: Erlbaum.
- Bower, G. H., & Mayer, J. D. (1986). *In search of mood dependent retrieval*. Unpublished paper, Stanford University, Stanford, CA.
- Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment*. New York: Basic Books.
- Butler, R. A. (1965). Investigative behavior. In A. M. Schrier, H. F. Harlow, & F. Stollintz (Eds.), *Behavior of nonhuman primates* (pp. 463-493). New York: Academic Press.
- Ekman, P., & Friesen, W. V. (1975). *Unmasking the face*. Englewood Cliffs, NJ: Prentice Hall.
- Fiske, S. T., & Linville, P. W. (1980). What does the schema concept buy us? *Personality and Social Psychology Bulletin*, 6, 543-557.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.
- Greenberg, L. (1986). Change process research. *Journal of Consulting and Clinical Psychology*, 54(2), 4-11.

- Greenberg, L. S. (1984). A task analysis of intrapersonal conflict resolution. In L. N. Rice & L. S. Greenberg (Eds.), *Patterns of change: Intensive analysis of psychotherapy process* (pp. 124–148). New York: Guilford Press.
- Greenberg, L. S., & Johnson, S. (1986a). Affect in marital therapy. *Journal of Marital and Family Therapy*, 12(1), 1–10.
- Greenberg, L. S., & Johnson, S. (1986b). Emotionally focused couples therapy: An integrated affective systemic approach. In N. S. Jacobson & A. S. Gurman (Eds.), *The clinical handbook of marital therapy* (pp. 253–276). New York: Guilford Press.
- Greenberg, L. S., & Safran, J. D. (1984a). Integrating affect and cognition: A perspective on therapeutic change. *Cognitive Therapy and Research*, 3, 559–578.
- Greenberg, L. S., & Safran, J. D. (1984b). Hot cognition: Emotion coming in from the cold. A reply to Rachman and Mahoney. *Cognitive Therapy and Research*, 8, 591–598.
- Greenberg, L. S., & Safran, J. D. (1987). *Emotion in psychotherapy: Affect and cognition in the process of change*. New York: Guilford Press.
- Izard, C. E. (1977). *Human emotions*. New York: Plenum Press.
- Kahn, S., & Greenberg, L. (1980). Beyond a cognitive-behavioral approach: Congruent assertion training. In C. S. Adamec (Ed.), *Sex roles: Origins, influences and implications for women* (pp. 127–138). Toronto: Eden Press.
- Lang, P. J. (1979). A bio-informational theory of emotional imagery. *Psychophysiology*, 16, 495–512.
- Lang, P. J. (1983). Cognition in emotion: Concept and action. In C. Izard, J. Kagan, & R. Zajonc (Eds.), *Emotion, cognition and behavior* (pp. 192–226). New York: Cambridge University Press.
- Lang, P. J. (1984). The cognitive psychophysiology of emotion: Fear and anxiety. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 247–272). Hillsdale, NJ: Erlbaum.
- Lazarus, R. (1984). On the primacy of cognition. *American Psychologist*, 39, 124–129.
- Leventhal, H. (1979). A perceptual-motor processing model of emotion. In P. Pliner, K. Blankstein, & I. M. Spigel (Eds.), *Perception of emotion in self and others* (Vol. 5, pp. 1–46). New York: Plenum Press.
- Leventhal, H. (1982). The integration of emotion and cognition: A view from the perceptual motor theory of emotion. In M. S. Clarke & S. T. Fiske (Eds.), *Affect and cognition: The 17th annual Carnegie symposium on cognition*. Hillsdale, NJ: Erlbaum.
- Leventhal, H. (1984). A perceptual-motor theory of emotion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 117–182). New York: Academic Press.
- Levey, A. B., & Martin, I. (1983). Cognitions, evaluations and conditioning: Roles of sequence and rules of consequence. *Advances in Behaviour Research and Therapy*, 4, 181–195.
- Martin, I., & Levey, A. B. (1978). Evaluative conditioning. *Advances in Behaviour Research and Therapy*, 1, 57–102.
- Perls, F., Hefferline, R., & Goodman, P. (1951). *Gestalt therapy*. New York: Dell.
- Plutchik, R. (1980). *Emotion: A psychoevolutionary synthesis*. New York: Harper & Row.
- Rachman, S. (1980). Emotional processing. *Behaviour Research and Therapy*, 18, 51–60.
- Rachman, S. (1981). The primacy of affect: Some theoretical implications. *Behavior, Research and Therapy*, 19, 279–290.
- Rice, L. N., & Greenberg, L. S. (1984). *Patterns of change: Intensive analysis of psychotherapeutic process*. New York: Guilford Press.
- Rice, L. N., & Saperia, E. (1984). A task analysis of the resolution of problematic reactions. In L. Rice & L. Greenberg (Eds.), *Patterns of change: Intensive analysis of psychotherapeutic process* (pp. 29–66). New York: Guilford Press.
- Safran, J. D., & Greenberg, L. S. (1982). Cognitive appraisal and reappraisal: Implications for clinical practice. *Cognitive Therapy and Research*, 6, 251–258.
- Safran, J. D., & Greenberg, L. S. (1986). Hot cognition and psychotherapy process: An information processing/ecological approach. In P. Kendall (Ed.), *Advances in cognitive-behavioural research and therapy* (Vol. 5, pp. 143–177). New York: Academic Press.
- Schacter, S., & Singer, J. E. (1962). Cognitive, social, and physiological determinants of state. *Psychological Review*, 69, 377–399.
- Tomkins, S. S. (1980). Affect as amplification: Some modifications in theory. In R. Plutchik & H. Kelerman (Eds.), *Emotion: Theory, research and experience* (Vol. 1). New York: Academic Press.
- Zajonc, R. B. (1984). On the primacy of affect. *American Psychologist*, 39, 117–123.

PART IV

**DESCRIPTION AND THE
ORGANIZATION OF BEHAVIOR**

CHAPTER 15

Prospects for Theoretical Progress in Behavior Therapy

R. S. Hallam

INTRODUCTION

Judging by recent texts on the subject and the range of issues discussed at international conferences, the growing points in the theoretical foundations of behavior therapy are to be found in all branches of experimental psychology. It would be presumptuous on my part to attempt to review these many advances. This book sets out to do just that. I intend instead to undertake a metatheoretical exercise. The successful expansion of techniques and theoretical models has now reached the point where behavior therapy is little more than a vague umbrella term. The purpose of this chapter is to analyze the conceptual foundations of the movement and to construct an argument for choosing concepts and assumptions that define a therapy as being of the behavioral type. The value of this exercise ultimately lies in its implications for future theoretical progress.

It is difficult, if not impossible, to talk about progress in the theoretical foundations of a movement that has not been satisfactorily defined, however provisionally. I will argue that in order to maintain a theoretical and practical focus on behavior in its current environmental context, concepts of meaningful action (acts) will need to be integrated into theories underlying behavior therapy. This may entail giving up assumptions that behaviorally oriented psychologists have held dear.

Definitions of behavior therapy like that of the Association for Advancement of Behavior Therapy (see Franks, 1984) are intended, it seems, to unify a heterogeneous group of practitioners under one banner. The argument goes that if this aim is achieved then little else matters apart from growth and manifest success of the movement. However, some commentators, like Franks (1982, 1984), see amidst this success signs of dissolution and fragmentation. This chapter is an attempt to respond to these concerns in a constructive manner but not in a way that can hope to please everybody.

THE CHANGING ROLE OF THEORY IN BEHAVIOR THERAPY

There are three main elements in the foundations of a recognizable form of psychological therapy (London, 1964). These are its value commitments (and implicit philosophical position), its theoretical base, and its techniques. Behavior therapy has undergone significant change over the years and the relationship between these three elements has been restated from time to time. A commentator who is perhaps best qualified to give a contemporary overview remarked that behavior therapy was undergoing an "identity crisis" (Franks, 1984). Many would agree that the three interweaving elements of behavior therapy no longer have the appearance, the resilience, or the interconnectedness of an identified school of therapy. One practical but important aspect of current concern is the image of behavior therapy and its success in the market place of competing professional groups and therapeutic schools. As London (1972) put it, therapists need theoretical principles to increase their confidence and to fight intellectual battles, noting that theories like psychoanalysis were "pretentious, respectable and smart." Barlow (1980) was concerned that behavior therapy would be more widely accepted by the public if it were presented without laboratory conditioning language.

However, aside from these questions of marketing the product, the main challenge to the movement appears to come from fragmentation within rather than from competition without. The importance of any specific framework (e.g., learning theory) has declined to the point that behavior therapy could best be described as the application of the methods and concepts of experimental psychology to human problems. Indeed, Yates, following M. B. Shapiro's model of the scientist-practitioner, advocated just such a definition in 1970. The concern of commentators like Franks (1982) is whether behavior therapy can survive intact without a theoretical core and a "behavioral model of Man."

There are two main trends in recent views on the role of theory in behavioral therapy. The first may be called technical and theoretical eclecticism. This sets out to develop an effective science-based technology originally focused on behavior though now more broadly aimed. Any theory will do as long as it is scientific and any technique is adopted as long as it can be objectively evaluated. The second major trend is to regard behavior therapy as *passé* (Lazarus, 1977), and subsume it within a theoretical framework that encompasses all forms of persuasion, control, or mere influence. This second trend, which reflects the wish to transcend behavior therapy, will not be discussed. A third possible trend, not much in evidence, is to redraw the boundaries around a distinctively behavioral form of therapy. This option will be discussed later.

TECHNICAL AND THEORETICAL ECLECTICISM

There is an obvious advantage in eclecticism. The eclectic therapist is less likely to miss out on exciting technical and theoretical ideas that may be found in the practices of competing schools or in the pages of unorthodox journals. Eclecticism could end up as an excuse for license but in the hands of a behavior therapist it is

supposedly tempered by “meticulous observation, careful testing of hypotheses, and continual self-correction on the basis of empirically-derived data” (Lazarus, 1977).

The abandonment by some workers of any rigid learning theory foundations of behavior therapy was understandable. These foundations were seen as a hindrance to the development of effective techniques and to an expansion into new fields of endeavor (London, 1972; Hersen, Eisler, & Miller, 1975). In part, the occasional denigration of learning theory can be attributed to the top-down hypothetico-deductive manner in which techniques were, at that time, rationalized theoretically. Empirical research into these techniques often failed to support their supposed theoretical foundations (e.g., aversion therapy, Hallam & Rachman, 1976). Understandably, behavior therapists began to research psychological problems in context and opened up an entirely new data base for theory. Furthermore, in the early seventies there was a tremendous expansion of interest in clinical problems among academic researchers (e.g., Seligman, 1975). Research findings from the animal learning laboratory were no longer assumed to apply automatically to human subjects as higher-order cognitive processes were needed to explain experimental results. Increasingly, the theoretical foundations of behavior therapy were based on studies of human learning.

The development of cognitivism in mainstream academic psychology initially posed a dilemma for the eclectic behavior therapist. At one and the same time, cognitive psychology was scientifically respectable but its methods and assumptions conflicted with the position taken by Skinner and like-minded behaviorists (see Skinner, 1985). This conflict has since resolved itself in the sense that cognitive and cognitive-behavioral techniques are seen by many as part of behavior therapy (e.g., see Bellack & Hersen, 1985). However, it is likely that theoretical disagreement will persist and that fragmentation will result. Genuine differences in the conceptual foundations of techniques are likely to lead to disunity and this cannot be avoided by incorporating a variety of techniques under an umbrella term.

The main disadvantages of eclecticism are those accompanying a lack of persistence in taking ideas to their limits. Longer-term and possibly slower development is forsaken for short-term goals. Behavior therapy would clearly have been the poorer had it not borrowed techniques from other therapeutic schools. However, in retrospect, one can see that the objective evaluation of techniques, however well done, does not necessarily advance the subject theoretically. A theoretical evaluation is essential if a technique is to be absorbed properly into the movement. With hindsight, we can see that London's influential polemic “The end of ideology in behavior modification” (London, 1972) was not entirely accurate in its predictions. In his view technology (and systematic practice) would eventually generate its own theory. He proposed an antitheoretical stance, maintaining that techniques in the social and behavioral sciences do not usually evolve out of theory. However, the history of research into various techniques of desensitization and exposure for reducing fears illustrates the point that practice needs theory as much as theory needs practice. An overemphasis on techniques led to a great deal of wasted research effort to study the effects of minor procedural variations. These often turned out to be theoretically uninformative because no theoretical questions were asked. Techniques were sometimes denigrated (e.g., systematic or imaginal desensitization) because better techniques came along (e.g., *in vivo* exposure). The failure to make a complementary

theoretical transition, however, has in turn diminished the value of studies on these techniques. In some cases, principles were discarded along with techniques because the techniques "failed to work." However, the principles may have been inappropriately applied (e.g., systematic desensitization to agoraphobia, Hallam, 1985b, p. 154). In any event, the present impetus for research into exposure techniques seems to have come from theoretical speculation (e.g., Bandura, 1977; Lang, 1979) rather than from systematic practice.

London (1972) did foresee the abandonment of "a premature and precocious search for general principles from which to get over-extended, or for a professional ideology to which to be committed." Learning theories and the data they generate continue to be a source of technical innovation but they are no longer the only theoretical base nor the primary ideological commitment of the behavior therapist. Franks (1984) sums up the sentiments of many behavior therapists today when he says that

the thinking behavioural clinician questions both practice and technique, entertains alternative hypotheses to explain clinical phenomena, and never accepts on blind faith the teaching of any authority or training institute.

The value commitment has essentially shifted from sets of assumptions about the learned nature of psychological problems (and hence learning theories) to certain scientific values and methodologies. Franks' characterization of the behavioral clinician does not always match up to reality but it may be sufficient to mark off the values of behavior therapists and distinguish them from therapists who seem dedicated to uphold the opinion of some one or other authority. But a questioning, empirical approach is not exclusive to behavior therapy and unless it is combined with a coherent theoretical framework, it is, as Franks (1982) suggests, a recipe for possible disintegration.

The eclectic approach I have just outlined has gone hand in hand with a view of behavior therapy as an applied natural science. I will now examine the implications of this view.

BEHAVIOR THERAPY AS APPLIED SCIENCE

There is much to be commended in a school of therapy that is theoretically grounded in a science of experimental psychology, wedded to empirical methods, and rigorously examines its own effectiveness (e.g., see Wilson, 1982). Behavior therapy has become an applied off-shoot of psychology as a natural science, that is, a psychology seeking universal principles based on natural observation. This is the received view of experimental psychology and it is often presented forcefully. Valentine (1982, p. 5) puts it this way.

Many psychologists have failed to realize that, as scientists, their concern must be with the generalities of behaviour. As previously noted, a prime characteristic of behaviour is modifiability. From this it follows that psychological science can never be concerned with the content of behaviour (because this must necessarily vary) but only with the principles of adaptation. This puts it squarely in the realm of biological science.

As Valentine views it, social science is “an entirely different pursuit.” Similarly, Kanfer and Hagerman (1985), in accordance with this view of psychology, portray behavior therapy (and modern psychiatry) as the product of the naturalizing approach that has rid therapy of its historical roots in philosophy and theology. Consistent with the natural science view, these authors argue that the information-processing paradigm can be integrated into the theoretical foundation of behavior therapy on the grounds that it is methodologically sound, deals with universal characteristics, and is physicalist.

I would like to argue that an understanding of universal psychological processes, though vital, cannot ever be sufficient as a theoretical basis of behavior therapy. Content as well as process enters into therapeutic practice in obvious ways, for example, the description of problems and the selection of targets. A knowledge of process helps to explain the form that content takes but this form is also determined by local and particular conditions that may only be comprehensible in historical terms.

A neglect of the social and historical significance of behavior is indicated by the terminology applied to the aims of therapy (goal, target, behavioral objective, problem, etc.). Although one can see how a content-free psychology is forced into this position, it should also be evident that even “meaningless” behaviors, such as obsessional rituals, are inappropriate forms of meaningful acts (checking, washing, counting, etc.). The social significance of the act includes a tacit understanding of how much checking/washing/counting is normal (the client may not share this) and a tacit understanding of what these acts achieve. These understandings of the problem are clearly relevant to the suggested goals and agreed objectives. In broader perspective, it is desirable that clients do not view their problem as an entirely individual matter having no connection with social conventions (the latter often being outmoded or self-contradictory). A shared discussion of social and political issues can, in this way, facilitate constructive solutions rather than leave the client with nothing but a self-ascription of abnormality.

The importance of a theoretical analysis of content has been overlooked because the general good sense of therapists and their knowledge of the culture has been taken for granted. Furthermore, it is widely assumed that clients know what they want or can be helped to discover this without imposing the therapist's values. As long as clients' goals are ethical and adaptive, the purpose of therapy, it is assumed, does not present theoretical problems. This outlook is compatible with the concept of the consumer exercising a free choice in the market place, selecting the type of service (therapy) and the best product in its range. This myth is sustained because it accords with the usual manner of offering personal services. A conceptual analysis of why it is that certain types of problem present themselves (e.g., unassertiveness in women) is thereby rendered superfluous or of peripheral interest, that is, the social context of particular problems is left unanalyzed. The only relevant theory seems to be that which underlies the process of change.

Adhering to the natural science view, behavior therapy becomes incapable of conceptualizing the content of human problems and through content the social milieu that generates that content. In an earlier and more behavioral period acts were generally defined in everyday terms (e.g., picking-up-cutlery,” Ayllon & Azrin, 1964).

Content was identified with commonly accepted and nontechnical meanings. More recently, increasing reliance has been placed on psychiatric descriptions of problem content and on intrapsychic (cognitive) structures. The focus has shifted from an analysis of behavior in specific contexts to clinical disorders and clinical populations. Conceived (wrongly) as natural entities, clinical disorders (i.e., the pathological conception of unwanted, deficient, or undesirable behavior) have become the staple diet of behavior therapists. This, in essence, is the way that many behavior therapists have come to solve the problem of content. Because disorders are regarded as natural entities, no further theorizing about the social origins of these "disorders" is required. It is wrongly believed that a certain pattern of natural processes underlies them.

Bandura, (1969, p. 111) recommended that intended goals be defined in terms of observable behavior and not hypothetical internal states. In 1984, Bandura regrets that "in virtually every respect psychological services mimic traditional medical practices" and that "the relevance of research, the adequacy of behavioral analyses, and the utility of psychological procedures all tend to be measured against the pathology metaphor." Psychiatry has come to be accepted as a complementary natural science. Clinical disorders are defined in terms of objective operations that can be precisely stipulated; however, such operations have as much in common with scientific method as accountancy. They are invented by committees and are admittedly based on personal opinion.

The pathology metaphor, transmitted culturally and internalized by the client, is presented back to the therapist as a disorder, thereby authenticating its supposed objectivity (Hallam, 1983, 1985b). Complaints are regarded as natural phenomena. In fact, there is no compelling reason to suppose that psychological complaints (content) correspond in any simple way with psychological processes and can be reduced to them. To take but one example, a model of the depressogenic process (Lewinsohn, Hoberman, Teri, & Hautzinger, 1985) has come to incorporate so many psychological processes and so many culture-specific concepts that we must begin to doubt the reasoning behind a reduction of the phenomenon to natural processes. Emotional states like depression, and also anger (Averill, 1983) and anxiety (Hallam, 1985b) should perhaps be viewed, first and foremost, as socially constituted, that is, as deriving meaning from a social/linguistic context and rule-governed social practices. There is now a gathering number of philosophers and psychologists (see Rorty, 1980) who regard references to emotions as intentional in Brentano's sense (Brentano, 1874/1973) and not descriptive of natural things or processes.

Recent analyses of the conceptual foundation of Psychology (see Harré, 1983; Gergen, 1982) highlight the need to examine the sociohistorical character of psychological knowledge but do not necessarily lead to a rejection (e.g., Winch, 1958) of the concept of a natural science of human behavior. Reasons can be regarded as a kind of cause (Locke & Pennington, 1982) even if the contents of reasons cannot be predicted on the basis of universal psychological processes. Behavior therapy, I suggest, is therefore in need of a dual foundation in natural and social science so that human problems are considered from both perspectives.

I turn now to consider the behavioral legacy of behavior therapy and how this might be developed to satisfy this aim.

WHAT IS BEHAVIORAL ABOUT BEHAVIOR THERAPY?

The question whether behavior therapy is behavioral must arise when one surveys the range of techniques that are included under this heading. The philosophical forms of behaviorism have largely been rejected and the methodological forms do not provide an unambiguous conceptual basis. Irwin (1978, p. 82) now considers it time "to get the behaviorism out of behavior therapy." What he seems to be recommending is pragmatic behaviorism, that is, therapy that aims to modify behavior in its situational context. But for Irwin, mentalistic explanations should not be rejected. According to this recommendation, behavior therapists would be distinguished by their preferred techniques and targets of intervention, that is, training new skills, habits, and other behaviors rather than, say, being most concerned with thoughts, memories, and physiological processes.

This pragmatic and technological stance does not supply a coherent meaning to the behavioral analysis of problems and it leads to theoretical eclecticism. I will first consider some objections to the more extreme behavioristic positions and suggest some alternatives to a pragmatic behaviorism that seems to have altogether too few conceptual constraints.

The insistence on measuring specific publicly observable behaviors, which was part of the positivist program in behavioral psychology, has paid off in a number of ways in the applied setting. Problems like alienation, which seemed to be purely subjective phenomena, were approached from quite a different angle. In sum, applied behavioral analysis was successful. However, as a philosophical position behaviorism has been rejected. Metaphysical behaviorism (that minds or mental states do not exist) was regarded as untenable, and analytic behaviorism (that all statements about the mental can be translated into statements about behavior or tendencies to perform behavior) was never widely endorsed. The weaker and more favored position is that an adequate psychology can be developed out of an analysis of behavior alone. Hypothetical constructs of various kinds have been postulated but these are tested with reference to observable behaviors, and purport to represent mechanisms that are neither mentalistic nor entirely neurophysiological.

As a basis for therapy, methodological behaviorism presents obvious difficulties. Therapists have to talk the language of mentalism in order to communicate with their clients. As it happens, mentalistic explanations do often suffice, because the therapist may not have any other conceptualization (behavioral therapy is inadequate). A behavioral analysis of a client's problems (often based on secondary diary data) is sometimes no more than a redescription of the terms of a mentalistic explanation tied in, here and there, with some public observations. When units of behavior are measured, the units are usually modeled on the preconstructed concepts of everyday language (e.g., smiling, running away). Similarly, situational events are described in terms that carry an implied meaning for the behavior being studied.

These are the realities of everyday clinical work. Even though a behavioral analysis of event contingencies is limited, it needs supplementing rather than replacing. The process of translation from everyday language to behavioral analysis entails the abstraction of functional classes of behavior, and meaningful content is ignored

by focusing only on general properties of stimulus cues, consequences, etc. In this way behavior therapy has rightly laid claim to being scientific in the sense that it adopted descriptions of behavior that have a potential universal application.

Given the difficulty of applying a behavioral analysis to complex human problems, there has naturally developed a tendency to water down the meaning of the term *behavioral*, and to produce explanations of a mentalistic or cognitive nature. These trends will now be discussed with a view to proposing a form of behaviorism that supplements rather than supplants behavioral analysis.

The positivist program in science included the supposition that public definitions and public tests of all scientific statements would allow the resolution of paradoxes and prevent future ideological misuse in the name of Science (see Flanagan, 1984, Chap. 4). When a concept was defined operationally its meaning was exhausted by the tangible physical operations used to decide whether or not to apply the concept. Anything one might want to say about the concept other than this was "surplus meaning." However, this strict approach to scientific method proved unworkable and all manner of hypothetical constructs are now found in behavioral psychology.

There are few behaviorists today who would want to limit their explanatory concepts to operationally defined (intervening) variables. A cogent behaviorist position is one in which explanatory constructs are invented to explain patterns in observations of *behavior in context*. Thus, a pharmacologist studying the site of action of a new drug might infer constructs from behavior but could not be said to be a behaviorist. There seems little sense in calling behavior therapy behavioral because it claims to be based on a scientific framework. Furthermore, attempts to define operationally the everyday (and mentalistic) concepts that are necessarily employed in therapeutic exchanges does not make them scientific or, in any sense, relate them to behavioral theory. For example, operational definitions of anxiety inevitably have surplus meaning because the original nonoperational meaning of the term is retained in some form. If this is denied, it is up to the scientist-practitioner to argue the case that this particular everyday concept corresponds to a useful scientific construct, for example, by showing in what way it would be more useful than a construct derived from, say, jealousy or alienation (Hallam, 1985a).

It seems that if we want to retain everyday constructs like anxiety (and hence their surplus meaning), then we cannot fully translate them into universal behavioral concepts because anxiety is not a natural category. Constructs like anxiety are part of a wider network of culture-specific practices some of the meaning of which is lost when a behavioral reduction is attempted.

MENTALISM AND COGNITIVISM: A BEHAVIORAL RESPONSE

A major dilemma for the behavior therapist is to explain and assess clients' problems when these are expressed in terms of mental states. By mentalism I mean the assertion that mental states exist independently of their manifestation in behavior, are not reducible to behavior, and can provide a causal explanation of behavior. Mentalistic explanations are culturally specific and in Western culture generally refer

to personal faculties and future goals. These explanations appear to be limited to acts rather than mere movements or reflexes.

One response to the dilemma is to reject behaviorism and take on board mentalistic explanations (e.g., Beck, 1984). Beck's approach is promising in terms of its efficacy, but apart from noting that its theoretical basis is not behavioral in the sense defined earlier, I will not discuss it further.

The solution of the cognitivist is to speculate more freely than the behaviorist about the mechanisms taking place within the organism. Like behavioral science, cognitive science seeks universal explanations. Kosslyn (1984) attempts to elucidate the differences between mentalistic (intentional) explanations and cognitive hypotheses. Whereas, he says, cognitive representations are taken actually to exist "in the head," "no intentional explanation can correspond in a simple way to a single internal event occurring in one person's mind." This may be taken to imply that statements of fact in cognitive science are reducible to physical phenomena in the brain. Kosslyn goes on to argue that the terms that make up mentalistic explanations depend for their meaning on the way they are used by a community of speakers. These meanings are negotiated, and may be revised on the basis of further information. A cognitive representation, Kosslyn says, is not something one can be right or wrong about, presumably (and paradoxically) because, in his view, it does not point to anything beyond itself.

Whether cognitive science has yet succeeded in its aim of becoming a natural science is hotly debated. Harré (1983, p. 10) charges cognitive psychology with transferring the truths of commonsense psychologies into the scientific mode by "transforming personal functions into mental organs (or in cybernetic terms, processing modules)" (see Harré for a further development of this theme).

Reda and Mahoney (1984) classify cognitive models as "surface structure associationist" or as models that direct attention to the "core ordering processes" of the human nervous system. The latter amounts to a biological conception of the origin of knowledge, in which knowledge structures are evolutionary patterns of information gathering and processing (Guidano, 1984). The aim of therapy is therefore seen as the modification of a client's conceptual frame for apprehending events. The biological conception of knowledge also leads to an affinity with the neuroscience aim of discovering a genetically programmed brain code. This could take psychological therapy into the arena of neurochemical intervention (Gray, 1985).

Arguments between cognitivists and behaviorists have been raging for several decades in behavior therapy journals (e.g., Breger & McGaugh, 1965) and there is no need to rehearse them (see Schwartz, 1982). The contrast I intend to draw between behavioral and cognitive therapy is a contrast at the theoretical level. The distinction can also be made, though less relevantly, in terms of the targets selected for intervention (e.g., changing beliefs rather than habits) or in terms of specific procedures (e.g., verbal persuasion versus skill training). These concrete distinctions are associated with theoretical concepts but do not clearly differentiate therapeutic schools. Behavior therapists are rightly concerned with problems presented as ways of feeling and thinking, and cognitive therapists do of course suggest behavioral exercises to change clients' conceptions of themselves.

A major difference between cognitive and behavioral theory is that cognitive structures and processes are conceptualized as properties of individuals (i.e., subjectively as modes of apprehension, etc.) whereas the significance of behavior is inferred in an objective fashion as functional elements in specific social and environmental contexts. How this significance is represented in the individual is of less interest than how it is publicly expressed as a network of influences between individuals and between individuals and the situations they find themselves in.

So, for example, maladaptive cognitions are seen as properties of persons, whereas maladaptive behaviors are more likely to be explained by situational and interpersonal processes. In clinical contexts, it is the individual who requests help and so cognitivism and individual therapy have a natural affinity. The individual internalizes social and environmental influences in order to act adaptively and so individual cognitive structures are a microcosmic reflection of the wider social order. It is therefore quite natural to study what an individual perceives and thinks rather than to study the influences these cognitive structures represent. Why *X* believes *Y* may be of less interest clinically than the fact that he or she believes it, that it is irrational, that it has certain consequences, and that it is open to influence. However, in addition to a proper theoretical concern with the cognitive processes of information transmission and change, it is also of theoretical interest to ask why *X* believes *Y* and why it has the consequences it does. If one does not accept the thesis that a theory of cognitive content is discoverable by studying the mind as a complex computing machine, these questions lead us directly into the interpersonal and public domain, that is, into the social construction of meaning.

What the behavioral approach has in common with social constructivism is a shared emphasis on behavior in the public domain. It might appear that a cognitive approach to the study of meaning is the obvious one to take. However, cognitive hypotheses, considered as statements about the natural world (and, of course, verified by public observations) are no closer to the elucidation of meaning than are naturalistic hypotheses about behavior.* In the cognitive approach, the contents or instruments of cognition (e.g., practical taxonomies, algorithms) that are developed in determinate historical conditions are elevated to the level of hypothetical universal processes. This is achieved by internal (structural, logical) analyses of meaningful content, which is reduced to finite sets of features, decision rules, and so forth. The latter are conceived as instructions to behave that bear little relationship to the communicative and practical situations in which meaningful content was originally (i.e., historically) developed.

An emphasis on meaningful behavior (acts or actions in situations) preserves the unity of psychological phenomena and the integrity of the actor. In other words, actors participate in situations (rather than respond as biological organisms) through activities that are required and maintained by situations and that culminate when the purpose of the act is fulfilled. Such activity need not require a prime mover in the form of an ego, mediating mechanism, or cognitive program. Moreover, insofar

*In what might be termed vulgar cognitivism, the weight of explanation is on concepts in everyday use. As noted earlier these are useful, even essential, in therapy, but they are not scientific concepts. They do not develop or advance in the way that formal knowledge advances.

as situations are organizing (contain sequences of instigations to act) the idea that actors represent all the features of the environment to which they respond seems unnecessary. To use an analogy, a lock can perform its function without a representation of the key—it need only respond to one key and not others. This mechanistic analogy should not be taken too literally but it is intended to make the point that actors inhabit the world and not, as some cognitive theorizing would have, representations of the world.

The argument I am making concerns levels of analysis and proposes a macroscopic level of analysis in which behavior is chunked into units that extend far beyond the usual time-limited definitions of a response. Sarbin (1985), for example, adopts the metaphors of drama and rhetoric to provide this macrostructure. That a sequence of behavior can be construed as an unfolding reflection of a dramatic plot, that is, as teleological, need not be regarded as antiscientific or preclude mechanistic analyses at a more microscopic level. The instigations to follow a plot are already present as antecedents in culture, myth, language, and various forms of social encouragement or inducement. And the outcome of a behavioral sequence of this type is not inevitable—accidents of life, situations presenting incompatible demands, and an intrinsic cultural diversity guarantee unexpected endings.

From a cognitive viewpoint, a plot or script exists in a person's mind, guiding or instructing behavior. The contrasting behavioral view is that plots/scripts are embodied in the real world, in its institutions, in social practices, and in the information media. Their embodiment within the individual, through acculturation, does not necessarily leave actors in a position to make an informed commentary on what is in their mind even though they may know what to do in a given context. A behavioral approach, in these circumstances, is to make inferences about the structure and structuration of behavior from a study of its antecedents and consequences, including actors' accounts of their actions. An alternative approach, tempting to the cognitivist, is to infer from accounts and behavioral performances an idealized, contradiction-free set of rules that simulate, as far as possible, the vagaries of the observations. Individual behavior is then regarded as an imperfect performance of this unwritten score.

It is easier to support a preference for the behavioral approach by example than by positive argument. So, for example, Lang (1983) has differentiated pathological groups on the basis of their typical affective memory networks. Lang has found differences in imagery ability of focal phobics and agoraphobics. It is inferred from these and other data that the latter group has less coherent affective memory networks and that the arousal response pattern has much broader associations. This conceptualization, even if broadly correct, focuses on specific characteristics of pathological individuals with the implication that the former are grounded in the properties of CNS structures. The fact that this state of affairs may be the product of an individual's life history and current existential choices is therefore lost from sight, even if it is not overtly denied. What may prove to be the result of poorly organized or socially maladapted behavior, namely high affective arousal, is taken instead as the point of clinical interest.

I suggest, therefore, that the implications of behavioral and cognitive approaches to therapy are different in important ways and that there are some reasons for

preferring the former. However, as behavioral processes are currently described they do not account for the content of human problems and it is to this aspect I briefly return.

BEHAVIORISM AND THE CONTENT OF BEHAVIOR

As a way of facilitating theoretical progress, I have argued that a distinction should be drawn between theories of universal psychological processes, and a psychological or sociological understanding of the content or meaning of what is organized. In this way the theoretical foundations of behavior therapy, which are made up largely of theories of behavior change, would be enriched by social science accounts of contemporary social behavior and its origins. Although there are examples of this kind of theoretical interchange (e.g., behavioral economics) there is little sign of it in the field of personal problems. The pathology metaphor has continued to hold ground whether presented physiologically, behaviorally, or cognitively.

One example to illustrate the relevance of a social perspective is Averill's work on anger and aggression (Averill, 1982, 1983). This could be described as an objective analysis of scripted behavior or as descriptive sociology. The script concept was suggested early on as an alternative to S-R terminology (Breger & McGaugh, 1965) and there have been, of course, more recent advocates (e.g., Goldfried & Padawar, 1982). However, a script need not be taken as a cognitive construct; as noted earlier, a behavioral script can be regarded as a descriptive unit applying to much longer time spans than is usual in behavioral analysis and normally as involving other persons. Scripts are intentional, that is, directed towards a completion point or payoff (e.g., courtship culminating in marriage).

Averill's data consisted of diary reports of anger episodes of normal individuals. Averill concluded from his analyses that anger is a highly interpersonal emotion that cannot be fully understood apart from the social context in which it occurs. The typical instigation to anger is a value judgment and, more than anything else, anger is an attribution of blame. It seems that contrary to some prevailing views, the consequences of an expression of anger are more likely to be described positively than negatively. Averill contrasts his approach with that of emotion theorists who have taken a purportedly scientific ahistorical analysis of antecedent causal events. He observes that the everyday phenomena with which we are often most concerned turn out to be teleological, historical, normative, and valuative. This type of analysis is surely necessary if we are to develop an adequate theoretical foundation for techniques for assertion training and anger control.

To conclude, in opposition to the subjectivist trend in theorizing, I have suggested that the theoretical constructs likely to be of value to a distinctively behavioral form of therapy should be inferred from observations of behavior and action. However, present concepts of behavior appear to be inadequate, especially with respect to long temporal sequences and interaction between behaviors. So, for the same reasons that the content of cognitions has acquired theoretical significance, it can be argued that behavior should be studied as meaningful sequences (acts, practices, etc.) in relation to their historical, moral, ideological, and institutional supports. Behavior therapy would thereby acquire a social and natural science foundation. In

addition to its valued role as a technology of behavioral change, it would also become informed in the craft of possible and creative change.

CONCLUSION

This chapter has discussed a number of interrelated themes in which the arguments converge on a preference (methodological and ideological) for inferring psychological constructs from behavior and its functional relationship to environmental events. It has been argued that there is some point in defining a form of therapy that is properly called behavioral in the sense that it has a behavior-theoretic foundation. Behavior therapy is contrasted here with theories that employ behavioral observations merely to test out hypotheses about internal/subjective (e.g., mental, cognitive, medical) constructs. In behavior theory, behavioral observations are the factual base from which constructs are inferred and subsequently tested. The efficacy of therapy based on nonbehavioral concepts is not of course disputed.

The chapter has further argued that although theories about universal psychological processes are essential as a theoretical base for behavior therapy, they are not in themselves sufficient to account for the historical and cultural forms in which problems are presented. It has been suggested that these historical forms are conditioned by linguistic and rule-governed social practices. It is maintained that the social scientific analyses to which these practices have already been submitted offer insights into the social instigation of maladaptive behaviors and that these behaviors should not be regarded simply as the expression of natural or biological disorders. Thus, to illustrate, therapy, which had as one of its aims the modification of a person's beliefs, would draw on theories about the social and historical origins of those beliefs as meaningful responses to a human situation as much as it would draw on cognitive theories about the nature of universal cognitive processes.

Behavioral therapists identify with an educational and not a medical model of intervention. This chapter has emphasized that reeducation takes place in a cultural, moral, and historical context. Therefore, behavior therapy cannot be reduced to a set of techniques for modifying a set of disorders without losing an important sense of the word *educational*.

ACKNOWLEDGMENTS

I am indebted to Mary Boyle and Simon Jakes for highly valued comments on a draft of this chapter. I would like to express thanks, also, to Glenice Gould for typing the manuscript.

REFERENCES

- Averill, J. R. (1982). *Anger and aggression: An essay on emotion*. New York: Springer-Verlag.
- Averill, J. R. (1983). Studies on anger and aggression: Implications for theories of emotion. *American Psychologist*, *38*, 1145–1160.

- Ayllon, T., & Azrin, N. H. (1964). Reinforcement and instructions with mental patients. *Journal of Experimental Analysis of Behavior*, 7, 327–331.
- Bandura, A. (1969). *Principles of behavior modification*. New York: Holt Rinehart & Winston.
- Bandura, A. (1977). Self efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191–215.
- Bandura, A. (1984). On paradigms and recycled ideologies. In S. Rachman (Ed.), *Contributions to medical psychology Vol. 3* (pp. 225–245). Oxford: Pergamon Press.
- Barlow, D. H. (1980). Behavior therapy: The next decade. *Behavior Therapy*, 11, 315–328.
- Beck, A. T. (1984). Cognitive therapy, behavior therapy, psychoanalysis and pharmacotherapy: The cognitive continuum. In J. B. W. Williams & R. L. Spitzer (Eds.), *Psychotherapy research: Where are we and where should we go* (pp. 114–135). New York: Guilford Press.
- Bellack, A. S., & Hersen, M. (Eds.). (1985). *Dictionary of behavior therapy techniques*. New York: Pergamon Press.
- Breger, L., & McGaugh, J. L. (1965). Critique and reformulation of “learning theory” approaches to psychotherapy and neurosis. *Psychological Bulletin*, 63, 338–358.
- Brentano, F. (1973). *Psychology from an empirical standpoint*. London: Routledge, Kegan & Paul. (Originally published 1874.)
- Flanagan, O. J. (1984). *The science of mind*. Cambridge, MA: M.I.T. Press.
- Franks, C. M. (1982). Behavior therapy: An overview. In C. M. Franks, G. T. Wilson, P. C. Kendall, & K. D. Brownell (Eds.), *Annual Review of behavior therapy, theory and practice, Vol. 8*. New York; Guilford Press.
- Franks, C. M. (Ed.). 1984. *New developments in behavior therapy: From research to clinical application*. New York: Haworth Press.
- Gergen, K. (1982). *Toward a transformation in social knowledge*. New York: Springer Verlag.
- Goldfried, M. R., & Padawar, W. (1982). Current status and future directions in psychotherapy. In M. R. Goldfried (Ed.), *Converging themes in psychotherapy: Trends in psychodynamic, humanistic and behavioral practice* (pp. 3–49). New York: Springer.
- Gray, J. (1985). A whole and its parts: Behavior, the brain, cognition, and emotion. *Bulletin of the British Psychological Society*, 38, 99–112.
- Guidano, V. F. (1984). A constructivist outline of cognitive processes. In M. A. Reda & M. J. Mahoney (Eds.), *Cognitive psychotherapies* (pp. 31–46). Cambridge, MA: Ballinger.
- Hallam, R. S. (1983). Agoraphobia: Deconstructing a clinical syndrome. *Bulletin of the British Psychological Society*, 36, 337–340.
- Hallam, R. S. (1985a). Anxiety and the brain: A reply to Gray. *Bulletin of the British Psychological Society*, 38, 217–219.
- Hallam, R. S. (1985b). *Anxiety: Psychological perspectives on panic and agoraphobia*. London: Academic Press.
- Hallam, R. S., & Rachman, S. (1976). Current status of aversion therapy. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification*. Vol. 2. (pp. 179–222) New York: Academic Press.
- Harré, R. (1983). *Personal being*. Oxford: Blackwell.
- Hersen, M., Eisler, R. M., & Miller, P. M. (1975). Historical perspectives in behavior modification: Introductory comments. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification*, Vol. 1, New York: Academic Press.
- Irwin, E. (1978). *Behavior therapy: Scientific, philosophical and moral foundations*. Cambridge: Cambridge University Press.
- Kanfer, F. H., & Hagerman, S. M. (1985). Behavior therapy and the information-processing paradigm. In S. Reiss & R. R. Bootzin (Eds.), *Theoretical issues in behavior therapy*. Orlando, FL: Academic Press.
- Kosslyn, S. M. (1984). Mental representation. In *Tutorials in learning and memory: Essays in honor of Gordon Bower*, Vol. 2 (pp. 91–118). San Francisco: W. K. Freeman.
- Lang, P. J. (1979). A bio-informational theory of emotion imagery. *Psychophysiology*, 16, 495–512.
- Lang, P. J. (1983, September). *The cognitive psychophysiology of emotion: Fear and anxiety*. Paper presented at the NIMH Conference on Anxiety and Anxiety Disorders. Tuxedo, NY.
- Lazarus, A. A. (1977). Has behavior therapy outlived its usefulness. *American Psychologist*, 32, 550–554.

- Lewinsohn, P. M., Hobermann, H., Teri, L., & Hautzinger, M. (1985). An integrative theory of depression. In S. Reiss & R. R. Bootzen (Eds.), *Theoretical issues in behavior therapy*. Orlando, FL: Academic Press.
- Locke, D., & Pennington, D. (1982) Reasons and other causes: Their role in attribution processes. *Journal of Personality and Social Psychology*, *42*, 212–223.
- London, P. (1964). *The modes and morals of psychotherapy*. New York: Holt, Rinehart & Winston.
- London, P. (1972). The end of ideology in behavior modification. *American Psychologist*, *27*, 913–920.
- Reda, M. A., & Mahoney, M. J. (Eds.). (1984). *Cognitive psychotherapies*. Cambridge, MA: Ballinger.
- Rorty, A. O. (1980). *Explaining emotions*. Berkeley, CA: University of California Press.
- Sarbin, T. R. (1985, June). *Emotions as situated actions*. Paper represented at the Conference on the Role of Emotions in Ideal Human Development, Heinz Werner Institute of Developmental Psychology, Clark University, Worcester, MA.
- Schwartz, R. M. (1982). Cognitive behavior modification: A conceptual review. *Clinical Psychology Review*, *2*, 267–296.
- Seligman, M. E. P. (1975). *Helplessness: On depression, development and death*. San Francisco: Freeman.
- Skinner, B. F. (1985). Cognitive science and behaviorism. *British Journal of Psychology*, *76*, 291–301.
- Valentine, E. R. (1982). *Conceptual issues in psychology*. London: Allen & Unwin.
- Wilson, G. T. (1982). Psychotherapy process and procedure: The behavioral mandate. *Behavior Therapy*, *13*, 291–312.
- Winch, P. (1958). *The idea of social science and its relation to philosophy*. London: Routledge Kegan & Paul.
- Yates, A. J. (1970). Misconceptions about behavior therapy: A rejoinder to Meyer. *Behavior Therapy*, *1*, 113–114.

CHAPTER 16

Behavioral Assessment A New Theoretical Foundation for Clinical Measurement and Evaluation

Ian M. Evans and Brett T. Litz

INTRODUCTION

With proper measurement being so important for the progress of any scientific discipline, it is difficult to imagine behavior therapy developing in the absence of conceptual advances in behavioral assessment. In fact, careful assessment has been so integrally related to treatment design and evaluation in the behavioral tradition that some commentators are now bemoaning what they see as the growing segregation of professional concern for measurement and for intervention. The extremely rapid spawning of monographs, textbooks, and even journals specializing in behavioral assessment does seem to confirm some separation of identities. Before evaluating the situation too negatively, however, it is worth considering whether there might not be important concepts in assessment that require specific and detailed empirical analysis. One such topic could even be the investigation of how clinicians do, or could, use assessment information. It is arguable that the lackluster state of traditional psychological testing is a consequence of increasing preoccupation with the instruments themselves and less attention to their purpose and use (Glaser, 1981). In this chapter, therefore, we will examine some promising conceptual and methodological developments in assessment, but attempt to keep them closely related to the functions of behavioral measurement in the clinical endeavor.

When behavior therapy first emerged, there was no really viable assessment technology upon which to draw. The traditional personality testing movement had long since surrendered any role in the design and evaluation of specific treatments. The primary focus of the more objective tests was on the assignment of individual cases to diagnostic categories. In addition to the assignments being unreliable and

the categories ambiguous, the given diagnoses had little or no impact on the design of psychological intervention and only minimal influence on psychiatric treatment or type of placement. Projective techniques were at least ostensibly oriented towards identifying the nature of those individual personality characteristics that might guide the formulation of psychodynamic therapy. However, their inherent subjectivity probably reduced their value even for psychoanalysis, and certainly created double jeopardy for any relevance to behavior therapy. Of the extant approaches to personality measurement, only Eysenck's theory (Eysenck & Rachman, 1965) had any logical connection to early behavior therapy, because the role of conditioning was so central. In principle, Eysenck's personality system could make predictions about the acquisition of abnormal behavior or the optimal arrangement of learning variables for the design of treatment. In practice, however, few systematic efforts were made along these lines and the behavior therapy movement began simply to build up its own traditions of assessment, in relative isolation from what had gone before.

Inevitably, with the close connection between learning theory and behavior therapy, the methodology that did evolve was similar to that used in the conduct of basic animal research. Obvious, highly face-valid clinical phenomena became the dependent variables, with simple, easily quantified aspects of behavior being monitored in order to demonstrate the effectiveness of the intervention. Thus the measures were typically of single responses, recorded as directly as possible, and alterations in some parameter of the response constituted the reported clinical change. For a while, most of the empirical efforts in assessment were expended on obtaining the truest measure of the phenomenon of interest. As for evaluation prior to the design of treatment, problem behaviors were expected to be salient and self-evident, and the alternative to diagnosis, it was hoped, would be to conduct a functional analysis of the conditions that might be controlling the behavior.

Within the past 10 years or so, a variety of concerns about this basic strategy have begun to emerge. One—a debate over clinical relevance—is that statistically significant alteration in the parameters of some target behavior might occur but still not meet criteria for the practical meaningfulness of the change. Another is that an erroneous or trivial behavioral element might be used as the outcome measure instead of the complex and important phenomenon itself. Even more troubling has been the recognition that as the client inevitably exhibits a variety of potential target behaviors, the selection of the most problematic or serious requires a judgment for which assessment processes have not yet been well designed. And, somewhat related, the concept of the diagnostic category, or at least the syndrome, might not be as irrelevant as early behavior therapists had claimed, if, in point of fact, knowledge of the syndrome suggests where certain defects might lie or what other factors might require remediation.

In the material to follow we will describe some of the recent empirical and theoretical developments related to these themes, and consider in particular the methodological issues surrounding the refinement of clinical measures. As already implied, any such presentation is hampered by the lack of agreement on the theoretical basis for developing behavioral assessment. Originally, various authors argued that although there were conceptual differences between behavioral assessment and traditional assessment (Goldfried & Kent, 1972), it would be desirable for behavioral

assessment to build on the accomplishments of the past, and not to abandon traditional approaches (e.g., Evans & Nelson, 1974), paying careful attention to reliability, validity, and so on (e.g., Cone, 1977). A particularly strong case for “intelligently assimilating” the traditional measurement standards, has been recently pleaded by Barrios and Hartmann (1986). There are rumblings, however, that opinion is shifting toward arguing that an entirely new conceptual basis needs to be found (Cone & Hoier, 1986); as we, too, will argue in the next section, the psychometric test tradition may have just too many anachronisms to be compatible with continued progress in clinical behavioral assessment.

BEHAVIORAL ASSESSMENT AND MEASUREMENT THEORY

Traditional personality and ability tests draw on a well-developed set of principles—psychometric theory—and concomitant rules for their administration, scoring, and interpretation. The behavioral perspective on these principles is, in our opinion, beginning to have novel implications for the development of more sophisticated measurement standards.

RELIABILITY

Obviously, any measurement procedure must yield the same information each time it is applied to the same unvarying phenomenon. However, as behavior by its very nature varies from occasion to occasion, establishing reliability by test–retest correlational procedures is not highly regarded in behavioral assessment. This is especially true because the traditional index is not based on exact item agreement, but on a large group of subjects maintaining their relative rankings over the two measurement occasions. Interest in agreement has emerged from naturalistic observation where the agreement between two or more direct observers ensures that behavioral events can be unambiguously categorized. Agreement is not solely an intrinsic property of the observation system, although if the code uses ambiguous or poorly defined categories, agreement will be difficult to achieve. Agreement represents the capacity of the observers to make the judgments about behavior that are called for by the category system being used. Simple contingency tables (House, 1980) allow one to separate agreement on occurrences and nonoccurrences and thus interpret the nature of the observer bias when disagreements do occur. More complex category systems may require more carefully trained observers. Fortunately it is now recognized that the degree of agreement required is a function of the decisions being called for by the measurement procedure (Yarrow & Waxler, 1979).

The implications of this area for clinical practice have not always been obvious, with some behavioral assessors disappointed that elaborate observational systems have not swept into regular clinical use. In a sense, however, it is the observer who is the instrument, not the coding system, and the observation literature provides valuable insights into improving the objectivity of naturalistic methods for gathering important clinical data. Thus phenomena such as observer drift, although useful for

improving observational methods in general, perhaps serve an equally valuable function by preparing clinicians for the common sources of error we will mention toward the end of the chapter. Given that observers can agree but still be incorrect, the emerging measurement standard is veridicality. Techniques to enhance the veridicality of observation include: (a) training observers to some prior criterion of mastery rather than just noting percentage agreement after the fact, (b) using a specialized, criterion observer, or (c) documenting a percentage of the phenomena of interest on video tape, and reviewing them until consensus is achieved.

Some information can only be obtained by relying on the client as the observer. This is because many of the behavioral phenomena of interest to clinicians are private events, that is to say, *both* potentially verifiable events that generally take place in private (e.g., a person's sexual behavior, solitary drinking, etc.) and covert, nonverifiable events, such as thoughts and feelings. In both cases, assessment is largely dependent on self-reported information. Verbal self-report of cognition is a complex topic that will be mentioned again later. The reliability of self-reported potentially verifiable events should be posed as a question of accuracy; however, when the corroborating evidence is another person's observation, agreement is again the most neutral descriptor. A good example is O'Farrell, Cutter, Bayog, Dentch, and Fortgang's (1984) study of the agreement (high, as it turned out) between alcoholic patients' own reconstructed reports of drinking binges and their spouses' estimates of their drinking behavior. Sobell and Sobell (1986) reviewed a number of studies confirming that, contrary to common clinical lore, the self-reports of drinking and related behaviors by clients with alcoholism tend to be accurate.

Although self-reported information has been corroborated in many studies, it does not seem reasonable to conclude that all individuals with alcoholism or other problem behaviors are inevitably accurate in their self-report. The very nature of most clinical work belies any such assumption. Nor would we try to establish the intrinsic accuracy of a self-report "instrument," because it could be so easily faked. Thus it would be helpful if future researchers were to delineate more precisely the conditions under which client's self-reports are least likely to be accurate (motivation to please, impress, or deceive), the best ways to obtain some corroborating evidence, and the types of inconsistencies in self-reported information that might help reveal (as would the lie scale in a personality questionnaire) which clients are poor self-observers. The literature on self-monitoring (Nelson, 1977) shows some of the progress we are making in this direction, because we can now make data-based recommendations to ensure the most accurate recording, such as training in the procedure, making the client aware that some verification will be attempted, reward accurately recorded data, and so on (see Barlow, Hayes, & Nelson, 1984).

VALIDITY

As behavioral assessment procedures attempt to make as little use of inference as possible, traditional validation methods have never been of major concern. However, as there is a tendency, even among behavioral researchers, to go beyond the behavior measured to the concept supposedly underlying it, construct validity issues still abound. For example, changes in autonomic nervous system responses, overt

avoidance behavior, and self-report of anxiety do not show simple concordance, partly because of measurement artifacts (Cone, 1979; Kaloupek & Levis, 1983) and partly because of the relationships between these three response “modes” being systemic (interactive) not colinear (Evans, 1986). But it is only when these different measures are conceptualized as measures of a unitary construct (anxiety), rather than representing directly the variations seen in different—though related—behavioral phenomena, that the lack of simple covariation among them would occasion any surprise.

Whenever a single dimension or property of behavior is measured, one runs the risk of not adequately representing phenomena that have multidimensional properties. Penile tumescence, for instance, is a phenomenon encompassing various hemodynamic changes resulting in alterations in penile volume and rigidity. Penile *circumference* (as measured by a mercury-in-rubber guage) is a convenient *index* of these changes, however circumference changes parallel the early part of the pattern of alterations constituting erection much more closely than they do the later changes (Farkas *et al.*, 1979). Behavioral assessment is beginning to realize that measures supposedly having linear properties do not have equal validity throughout their range. Many important clinical constructs represent specific states of some variable phenomenon. This can be illustrated by penile tumescence again. Linear increases in circumference (or volume, or whatever) might represent a useful continuous dependent variable for laboratory research in sexual arousal. Impotence, however, the dimension of clinical interest, can not be defined by minor variations in volume, but according to absolute or threshold levels of rigidity sufficient (and prolonged enough) to effect intercourse.

Following similar reasoning, albeit in a very different context, Evans and his colleagues (e.g., Evans, Brown, Weed, Spry, & Owen, 1987) have developed an outcome measure of basic life skills that is oriented toward mastery of complex skill sequences (tasks) or “routines.” Mastery is judged by whether the routine achieves its critical effect, so that an individual with a disability can either do the task or not. No credit is given for partial accomplishments, because unless the function is successfully achieved, by whatever socially acceptable form, there is no obvious value in being able to perform elements of the routine. More typical measures of adaptive behavior allow the client being evaluated to earn a score by passing test items, no one of which represents successful attainment of a complete skill. Such measures cannot be thought of as truly criterion referenced, although their scores could have predictive capability if the instruments had been validated on a group of subjects whose daily living skills are known. Here again the predictive correlation might be high but the criterion variable—say, holding a job—is generally an all or nothing event in the real world. An ex-patient who almost has a job is unemployed.

Somewhat complementary ideas are beginning to pervade the more conventional behavior therapy outcome literature. In the past, the tradition of experimental research has resulted in outcomes of therapy studies being evaluated by statistically significant behavioral changes from preintervention levels, or by comparison to some untreated control group. Such changes in the parameters of the dependent variable could have theoretical significance (confirming a hypothesized relationship between independent and dependent variables), but are largely worthless for the evaluation of an intervention strategy. Clinical significance has been touted as of equal necessity,

but until recently there was really no attempt to define what that might be. Now, however, two approaches seem promising.

One of these is the more exact specification of what the acceptable or normal state of the phenomenon might look like. In the case of marital therapy, for example, Jacobson, Follette, and Revenstorf (1984) suggested that the criterion of success is whether the clients have moved from a dysfunctional to a functional range. In commenting on these authors' work, Wampold and Jenson (1986) reiterate the assumption that many clinical disorders are not dichotomous with normality. That may be the standard behavioral perspective (although one could hardly tell from the readiness with which contemporary clinicians refer to "phobics," "depressives," or "bulimics") but it could be that our measures have been too simple to capture the change in status when everyday behavior becomes maladaptive. Binge eating, for example, may phenomenologically be similar across patients and nonpatients, but its function (socially or hunger motivated, versus serving to reduce distress), temporal distribution (occasional rather than regular), and subjective properties are very different. Corrigan (1985) described one patient who self-reported an incident of bingeing after eating only three grapes for one meal. This person's subjective criteria obviously included dimensions such as compulsivity and inappropriateness of the food content for the meal (breakfast), in addition to simple volume or number of calories.

A second approach, similar to that of Jacobson and his colleagues, is also still in its infancy—template matching (Cone & Hoier, 1986). In this strategy the desired outcome of the intervention is determined (for content and level) by the characteristics revealed by a selected subgroup of comparison individuals. Essentially this means there is a fluctuating criterion in a criterion-referenced measure. The advantage is that the criterion group is not some averaged national sample, but the specific individuals with whom the client must interact and gain acceptance. A somewhat comparable idea has been articulated by Brown and his associates (Brown, Nietupski, & Hamre-Nietupski, 1976) for setting skill priorities in teaching persons with severe handicaps. The skill goals for such individuals, it is argued, should be determined by what is needed in the next environment to be experienced; they refer to this as the criterion of ultimate functioning.

IMPROVING THE VERIDICALITY OF MEASUREMENT

From the previous section it can be concluded that contemporary behavioral assessment theory is creating new standards and attitudes regarding what is important in psychological measurement. When measures are very direct and low inference, they should be accurate (veridical), ecologically relevant to the phenomenon of interest (valid), and contribute differentially to effective clinical decision making. We will discuss this third standard, utility, in greater detail later, and have already elaborated on validity criteria. To gauge just how far behavioral assessment has progressed in bettering the veridicality of its measurement procedures we will briefly mention some of the improvements in technique that have been emerging over the past few years. Rather than summarizing these by clinical phenomena (e.g., anxiety, depression,

social isolation, etc.), we will list a few of the more promising characteristics of behavioral measurement.

Typical versus Possible. One of the most valuable measurement principles to arise from the behavioral perspective is the difference between performance under optimal circumstances and under typical circumstances. The distinction is particularly well drawn in the literature on social skill assessment where it is now well recognized that how people act in a simulated assessment situation may be very different from how they would typically act in their everyday social environments (cf. Strosahl & Linehan, 1986). There would be little point in trying to validate a role-playing test with unobtrusive observation in the natural environment because the research on assertive behavior shows that subjective evaluation of the consequences influences such behavior (Kuppermink & Heimberg, 1983) and the nature of the consequences in the test situation and the real world are obvious. Conversely, of course, the determination that a client is capable, under some conditions, of producing a skillful social response, shows that the therapeutic focus may need to be on anxiety, interpretation of cues, and so forth, rather than on the content of the behavior.

A closely related point emerges from any type of test administered to clients whose motivation to perform or whose understanding of your expectations is suspect. If an autistic or mentally retarded child fails to perform a specific item on a standardized test, one cannot conclude that he or she could not perform the item, only that he or she did not do so. Providing reinforcement, improving comprehension of the instructions, and other adaptations not generally permitted with standardized instruments, generate the truly useful information about the factors that will determine an individual's performance.

Instrumentation. Instrumentation does not necessarily enhance veridicality of specific measures but does make possible an array of new ones. Fortunately, with its continued ties to general experimental and physiological psychology, behavioral assessment is potentially able to draw on many developments in basic psychological measurement. This is a period of great technological sophistication and so it is particularly ironic that some specialists in applied psychological measurement are still happily using colored ink blots or apparatus first devised by Binet and others at the turn of the century. Insofar as electronic or computer controlled apparatus is used in clinical assessment for measuring arousal, emotion, activity, force, and so on, it is mostly within the rubric of behavioral assessment that such advances have been made (see Rugh, Gable, & Lemke, 1986, for a recent review). Of course it takes time and effort and money to introduce technological advances, so that only methods with high treatment validity are likely to show much progress. Except for professional time, projective testing costs little, so the value of the resultant product need only be quite modest in order to perpetuate its use.

Multiple Response Monitoring. Behavioral assessment has made as yet unappreciated progress in distinguishing between different measures of the same construct (which should generally show some degree of concordance) and concurrent measurement of different aspects of behavior, which then provides new detail and information about clinical phenomena. Attitudes toward other children with handicaps, number of positive peer interactions, ratings of peers on a sociometric measure, and

counting the number of party invitations received, are all independent (though certainly not unrelated) facets of behavior, not signs of the hypothetical construct of friendship (see Evans, 1986, for a more extensive discussion). But by measuring these different behaviors one is likely to end up with a more precise picture of the general phenomenon, just as we do when measuring three parameters of the same specific behavior, such as the frequency, duration, and force of a self-injurious behavior in a person who is severely handicapped. Some behavioral assessment researchers are still likely to trot out ideas like convergent and discriminant validity, without fully recognizing the new concepts that emerge from thinking of the content and method of measurement (Cone, 1979). For clinicians this is not some esoteric argument; we benefit greatly from multiple measures to corroborate evidence and thus ensure that what has changed is not some arbitrary and perhaps uninteresting dimension of behavior.

Self-Report of Unverifiable Events. We have already commented on self-report of one's own behavior, which represents a special type of observer agreement issue. Behavioral assessment has also provided some clarification of the self-expression of feelings, attitudes, beliefs, and so on. Self-reports of cognitive processes are not likely to be useful, but self-disclosure of the content of one's thoughts (thinking out loud) could provide valuable information regarding thinking style as the clinician is then observing the content of thought, through speech, much as one would observe overt behavior, and with the same issues of sampling, reactivity, and other sources of bias.

Observation of Interaction. Direct behavioral observations has been the mainstay of behavioral measurement methods. It is certainly true that the behavior therapy research literature relies heavily on observational methods, so that standards for determining observer agreement, for sampling from the stream of behavior, and for defining behavioral categories have become quite sophisticated. In fact, good observation goes beyond the information that would be known simply through immediate sense data. Observation reveals new phenomena, one of which arises from the analysis of interaction—between parent and child, between married couples, between peers—being the relevant dimension. These new phenomena could not be discerned from considering one individual alone.

This is an exciting realization because it says much for the future viability of behavioral assessment, compared to assessment traditions focused exclusively on individuals and not their interactions with others or with their physical world (McReynolds, 1979). Important clinical phenomena can be defined only in terms of interaction (for example in distressed couples, unsatisfactory proposals are met with by counterproposals, rather than contracting—Gottman, 1979) and of dynamic progression over time. Gottman (1983), in his studies of how children become friends, provided an interesting example of the latter. A key process is escalation of “common-ground” activity, such as coloring side by side. In that social context escalation might involve suggestions of joint activities (“let's both color this blue”) or new rules (“we can't mess this up, OK?”). If the escalation is unsuccessful and a potential conflict brews, there is rapid deescalation, back to some common ground. It is difficult to imagine discovering meaningful social interaction deficits by individual assessment of a young child whose mother brings him to a clinical setting because he has difficulty making friends with other children in the neighborhood.

We can summarize this section by observing that behavioral assessment methods are still close to the research measures of the experimental laboratory. Traditional tests, with the exception perhaps of the ubiquitous self-rating questionnaire that has dominated so much of personality research, were developed specifically in clinical settings for applied problems. It is small wonder that they have ready acceptance in mental health and educational fields. Behavioral measures, derived from more basic behavioral research, are often cumbersome, time consuming, and generally not well suited for use with the individual patient in the office or clinic. To some extent complex measures can be simplified, and as more basic discoveries are revealed by intensive measurement they become the accepted clinical phenomena to be looked for. The clinician might well be able to discover, just by listening to and observing a couple in therapy, examples of counterproposal making rather than contracting, without having to replicate the research program that identified the importance of this dimension. Although it would be a pity to compromise measurement too far—there is important information that simply cannot be obtained in the clinician's office—another solution is to develop higher inference, indirect measures (Burns, 1980), but keep them more logically tied to the target behaviors of clinical interest. To illustrate this we need to consider the role of personality constructs in behavioral assessment.

COGNITIVE PERSONALITY ASSESSMENT

It has always been the hope of those trying to keep clinical practice within the mainstream of psychology as an empirical behavioral science that personality theory and research would provide a conceptual basis for clinical assessment. Certainly the search for generalized, temporally stable behavioral dispositions or traits that could offer a description of the individual in relation to some nomothetic dimension was the zeitgeist in clinical assessment for much of its history. Mischel (1968) was really the first to articulate the various misuses of trait-oriented assessment, basically showing that when clinicians used their favorite personality dimensions, their depictions of the clients' dispositions would rarely correspond with actual criterion behaviors in various life circumstances. Behavior was shown to be intimately tied to situational variables; assessment of performance in a specific situation defined by environmental contingencies would provide a more useful description of a client's problems as well as what had to be altered clinically. This challenge stirred an exciting debate among personality theorists, some of whom were anxious to find the coherence and stability in persons that they felt Mischel had eliminated in his writings (see Epstein & O'Brien, 1985).

A widely accepted resolution of the issue lies in recognizing the importance of *both* the situation and certain person variables; actually, a careful reading of Mischel (1968, 1973) would lead one to such a conclusion. Interestingly, clinicians and researchers have selectively abstracted what they perceived to be the essential aspects of the conflict, some seeing the establishment of broad person variables as paramount (e.g., Epstein, 1979), others seeing the measurement of the environment as the key to predicting behavior (e.g., Moos, 1973). Behavior therapists have prided themselves

in being on the right side of the debate; after all, they had been dealing with behavior as it unfolded in its natural context all along. However, an overly strict interpretation of the situational specificity of behavior would render most of the actual assessment behavior of behavior therapists invalid. Interactionism resolves that dilemma: behavior is a product of the situation and what the person brings to it. Basic behavioral repertoires (Staats, 1986) can mediate overt responding to the environment, and provide a foundation for stable, often generalized pattern of behavior. Thus the repertoire of more basic cognitive behaviors will yield the most clinically useful assessment information.

A rich source of information on the nature of basic cognitive repertoires comes from the growing rapprochement between social, personality, and cognitive psychological research (see Cantor & Khilstrom, 1982). The relevant conceptualization for cognitive-personality assessment is that people "respond flexibly to situations, as they construct them cognitively, and that they act behaviorally to transform situations so that they correspond more closely to their expectations" (Cantor & Khilstrom, 1982). Such a theoretical perspective provides the basis for an idiographic assessment of a person's specific, idiosyncratic responses to environmental cues that then mediate their behavior. This is neither a search for trait dispositions nor a strict, simplistic adherence to environmental contingencies.

Mischel (1973) postulated one of the more comprehensive theoretical schemes for specifying the cognitive-personality domains of greatest relevance to behavioral assessment. He distinguished five categories of person variables that represent the cognitive processes that are said to interact with situational information to account for the consistency of behavior as well as the intransigence of many clinical problems. These categories are construction competencies (an individual's knowledge, capabilities, problem-solving abilities); encoding strategies and constructions (how information is organized, abstracted, categorized, utilized, or an individual's implicit personality theories); expectancies (outcome and efficacy expectations); values and preferences (ratings of the affective valence of key reinforcing or aversive stimuli); and self-regulatory systems and plans (self-statements, goals, selective attention to salient information, memorial retrieval of clinically relevant information, metacognition, i.e., an awareness of one's own cognitive processes).

Perhaps the essential assessment issue arising from this general approach is the recognition that individuals can often cognitively transform their environment; they go beyond (adding to or abstracting out) the information presented in a stimulus array (see Alba & Hasher 1983, Taylor & Crocker, 1982). Individuals are aided in their perceptions of events by organized knowledge (schemata, scripts, frames) that may serve to fill in missing or essential pieces of information by providing default or hypothesis-relevant cognitions. Behavior is influenced (or biased) by such organized knowledge because an individual will respond in a manner that is consistent with preexisting notions of "how things work" in their social world. (Presumably this is true for clinical psychologists as well, thus permitting valuable insights into assessment methods, as we will discuss presently.) Such constructive processes are not new to psychology (see Bartlett, 1932), but they do represent a new challenge to the understanding of clinical phenomena (Hollon & Kris, 1984), as well as the assessment field (e.g., Merluzzi, Glass, & Genest, 1981). However, we need to add that although

we see these efforts as essential, they are still in their infancy in regard to practical clinical utility, or tying constructs like self-talk and memory retrieval to criterion, problem behaviors. The most challenging finding in recent years comes from Cutrona, Russell, and Jones's (1985) demonstration that cognitive events are not traits either: people do not have "attributional styles" but make different causal attributions about different life events.

PERSONALITY AND THE PURPOSE OF ASSESSMENT

We began this chapter with a consideration of the veridicality of information, and the discussion of personality followed from questioning the type of information clinicians most need. Before leaving the specific topic of personality, therefore, it is worth reconsidering the various ways in which personality constructs could actually influence the clinical endeavor. So far our discussion of personality is synonymous with understanding the commonalities of human behavior, that is to say, a technical equivalent of the intuitive knowledge we all have that allows us to predict how other humans may feel or act in a variety of situations. Clearly, this technical knowledge is what we are striving for when we talk of understanding humans in general and clinical patients in particular. But there are a few other, more limited ways that personality constructs might aid the design of effective intervention. One of these is the selection, for therapeutic attention, of the real issue, or, in behavioral terms, the behaviors (including cognitions) that are the most relevant independent (causal) variables for the most obvious presenting complaint (the behavior that has brought the person to therapeutic attention in the first place).

In behavioral assessment this search for the most mutable and influential element of the individual's repertoire is described as target behavior selection and will be discussed in the next section. For now, however, it should be noted that the search for relevant behavior as independent variables is a direct derivative of personality psychology and it demands some understanding of the organization of individual repertoires (Evans, Meyer, Kurkjian, & Kishi, in press; Voeltz & Evans, 1982). Behavioral consistency is deduced from the general rule that one behavior could be related to another in a predictable manner, and that assumption in no way precludes the certainty that this second behavior is also regulated by unique environmental independent variables and that even that relationship is mediated by other aspects of the person repertoire in the manner just described under cognitive variables. To help the reader follow this very abstract summary of a complex issue we can give a stylized example: the child whose frequent aggressive behavior (the complaint) is "caused" by other more basic behaviors (such as poor coping strategies for dealing with frustrations), although environmental events regulate and may have shaped the aggressive behavior itself (hitting being reinforced by parental acceptance) and cognitive behaviors (the interpretation of peers' actions as having hostile intent) mediate what would in other circumstances be neutral environmental stimuli (good-natured teasing by others). Incidentally, all these response relationships have been documented, separately, in research on child aggression (Crowell, Evans, & O'Donnell, 1987).

A second potential purpose of personality assessment is to provide the clinician with guidance regarding the most effective way to design treatment, not in terms of the focus, but in terms of its exact structure. At the most superficial level, this strategy is seen in such assessments as a reinforcement survey to select the most potent reinforcer for treatment, or the identification of a patient's interest, beliefs, and cultural values so that therapeutic suggestions may be focused in a mode that is acceptable and plausible. Although referring to this as superficial, we mean that theoretically—such assessments could be very important in making the difference between success and failure in the design of behavior therapy. Simple protocols that assess patient's factual knowledge about topics such as nutrition (in eating disorders) and sexual physiology and anatomy (in sexual dysfunctions) are extremely useful. We might also include in this category the gathering of information on a client's network of social supports to see who might be mobilized to assist in behavior change. Personality theorists who are trying to specify the match between persons and environments (e.g., Bem & Funder (1978) provide the theoretical foundation for describing situational templates, whose virtues have already argued.

The most sophisticated level of personality theorizing in clinical practice occurs when the style or parameters of intervention are chosen according to hypotheses regarding exact therapeutic mechanisms. For instance if an intervention (such as the bell and pad for enuresis) works by virtue of classical conditioning, and if different personality types (such as introverts and extraverts) differ in their conditionability, then one might be able to design treatment parameters (such as number of trials, intensity of the CS and UCS) according to the personality characteristics of the patient. Although there are the number of cases that could be cited where therapy has been designed along these lines (Eysenck, 1982) the approach still remains a hypothetical ideal—there are a lot of major “ifs” in the preceding statement of the strategy, despite Eysenck's theoretical approach to personality being couched in the same principles of learning and conditioning as behavior therapy. The lack of precision in either the theory or the behavioral techniques should not detract from pursuit of this model as an ideal for the future. Somewhat analogous arguments have been made about teaching-based interventions where in principle teaching style could be modeled to learner characteristics.

TARGET BEHAVIOR SELECTION

In the previous section we began to discuss the topic of target behavior selection. Recall the hypothetical analysis of aggression in a child. If the analyses presented were correct, what should be the target of the therapeutic intervention? Should it be the lack of more adaptive social skills for dealing with interpersonal conflicts, should it be the parental reinforcement for overt acts of physical aggression, or should it be the apparently irrational cognitive interpretations of other children's signals? Behavior therapy has approached the problem all three ways so a clinician might be justified in directing treatment to as many facets of response organization as possible. Ideally, however, assessment should be able to ascertain which of a wide variety of possible intervention targets would be the most productive. That, in a nutshell, is the issue

of target behavior selection, which emerges once it is accepted that what seems like the obvious problem (the complaint) is merely the most salient element of a repertoire having a wide variety of possible anomalies and dysfunctions.

As argued before (Evans & Wilson, 1983), there used to be a tendency for behavioral assessment to conceptualize target behavior selection as a choice among a set of equally evident and independent behaviors. In other words the process was that of identifying all abnormal behaviors or complaints, making a list of them, and picking one to modify, then moving on to the next, and so on until one had successfully intervened with everything on the list. Clinicians were given criteria to help them prioritize the order of intervention. Research methods, particularly single-subject designs derived from operant conditioning, encouraged this type of thinking: one response had to be selected at a time, the response selected for monitoring in the baseline was the same as the response being treated, and even if more than one response was formally measured concurrently, the resulting “multiple” baseline design was predicated on the assumption that there would be no response–response interactions between the two or three behaviors being measured.

This type of thinking has been criticized from a number of angles, such as the limitations of these research designs (Voeltz & Evans, 1983), the inevitable interrelationships among behaviors (Voeltz & Evans, 1982), and the inadequacy of the model that equated treatment focus with the prioritization of a set of potentially equal and independent target behaviors (Evans, 1985). Nevertheless, the type of systemic model building of how behavior within repertoires interact that Evans has advocated as an alternative conceptual metaphor has to be taken many stages further before a really viable assessment model evolves. For example, once we accept that clinical problems do not exist as isolated, single responses or behaviors, we then have a complex taxonomical problem.

To illustrate this, imagine a child client who exhibits the following problem behaviors: enuresis, fear of the dark, fear of dogs, and a below average reading ability despite normal IQ, and withdrawn, shy behavior in the classroom. First, it has to be recognized that not all these problems would necessarily have been identified at the same time or with equal facility; perhaps the parents brought the child complaining about the bed wetting, subsequent interviews unearthed the fear of the dark and of dogs, the reading difficulty was discovered only by norm-referenced testing, and the shy behavior only through interviewing the teacher. Second, it is clear that the problems listed are not at comparable levels, descriptively. And third, it is apparent that the various behaviors could be classified in a variety of different ways, for instance school problems (reading difficulty, shyness), manifestations of fear or anxiety (dogs, dark, peers), causally interacting behaviors (fear of dark contributes to enuresis), narrow-band syndromes (phobia), or broad-band syndromes (e.g., inhibitory behavior—enuresis, shyness, fear of dogs and dark, all forming a commonly observed cluster).

Thus there is both a discovery component and a classification component that is involved before one even gets to prioritization. There are criteria that can be applied when trying to organize these different behaviors, and these criteria fall into two groups—one related to outcome and one related to process. Outcome criteria are relatively straightforward: (a) Which behavior is most serious for the client and results

in the most distress to the individual or relevant caregiver? (b) Which behavior might have the greatest consequences for others, such as dangerous behaviors? (c) Which behavior has the greatest impact on future opportunities, such as educational placement or job possibilities? (d) Which behavior if not altered could get worse in terms of its negative impact for the client? Process criteria are a little more complex: (e) Which behavior is most amenable to change? (f) Which behavior, either of the identified ones or some other more basic one still unidentified, would have the biggest impact on the total repertoire if modified? (g) Which behavior if not specifically addressed might be successfully modified by natural contingencies over time? At last part of assessment research should be focused on gathering objective data that can help clinicians answer those types of questions.

CLINICAL DECISION MAKING

It is clear that behavioral assessment is now reorienting itself away from being a measurement technology toward examining the decision making needed for effective intervention. Meehl (1954, 1960) first expounded the view that it is untenable to assume that clinical psychologists' judgment processes are somehow immune from typical human cognitive biases and inaccuracies. His 1954 book sparked a great interest in the validity of diagnostic judgments, and a polemic over the relative merits of psychometric/statistical versus clinical/intuitive/inferential decision making. The assessment information of that time consisted mostly of projective tests, questionnaires, or interview data, and the metric was some known criterion—a categorical, discrete, or “bound” piece of datum (e.g., a diagnostic label). Because behavioral assessment is founded on a commitment to avoid such categorical labeling, the theoretical arguments and empirical data from this controversy have limited value. Thus we need to take a fresh and critical look at the cognitive and metacognitive activities that make up behavioral assessment judgments.

Some behavioral assessors have begun to examine the formal and informal rules of behavioral assessment as a process; Kanfer's work is especially important in this respect (e.g., Kanfer, 1985). There are important descriptions of the kind of information that must be gathered (e.g., Herbert, 1981) and the central value of hypothesis testing and the functional analysis as an assessment model (Nelson & Hayes, 1986). With the exception of a few specific studies of behavioral assessment decision making (e.g., Felton & Nelson, 1984; Hay, Hay, Angle, & Nelson, 1979; Wilson & Evans, 1983) we have no real knowledge of whether behavioral assessment techniques protect clinicians from judgment errors or merely add a new set of assumptions that bias the process. However, as there has been much recent interest in the behavioral literature regarding these biases, we will briefly review them. One essential caveat is worth stressing. The decision-making research has almost exclusively investigated the judgment behavior of individuals (often university students) who are required to make bounded, categorical decisions. The ecological validity of this information for clinical situations with experienced professionals is unknown.

The common biases that have been identified that may be of relevance for conceptualizing behavioral assessment as a protective strategy for error are confirmation biases, overconfidence, and the illusion of validity. It may prove useful to think of clinicians as possessing organized prior bits of knowledge or schemata that facilitate their clinical decision making (cf. Hollon & Kriss, 1984; Turk & Salovey, 1985). It is indeed adaptive to use such stored information in order to be able to select our essential bits of useful information amid a morass of potentially relevant cues or signals. Professional training creates schemata or implicit assumptions which can guide the kinds of questions that are asked in an interview, the types of assessment information sought, and the kinds of hypotheses engendered. Schemata have been found to provide "default values" under conditions of information uncertainty (cf. Hastie, 1981). This may be seen clinically, for example, when a therapist is confronted with inaccurate or incomplete information about a client's problems; prior knowledge or expectations drawn from psychopathology research, for instance, provide "best guesses," which can foster the provision of more complete information. Such an exchange of expectations might also be thought of as the cognitive basis of empathy, that is, knowing how a person may behave, feel, or think given scant, but key pieces of information.

Besides serving such an essential adaptive function, schemata can be utilized clinically in an inappropriate fashion, and can lead to biased information processing; our assumptions about missing pieces of (assessment) information may be invalid. We will seek information that will be congruent with a hypothesized schema, or perhaps not attend to information that is hypothesis irrelevant. Such *confirmation biases* can be seen as one of the key vagueries of clinical decision making. Unfortunately, there are few systematic data on the influence of such biases (or processes), clinically. In medical decision making, physicians have been shown to use such implicit (schematic) information processing, with some resultant biases: information that is hypothesis (diagnosis) irrelevant (e.g., inconsistent findings) is often disregarded, or favorite interpretations of data are overutilized at the expense of alternative hypotheses (Elstein, Shulman, & Sprafka, 1978).

Research in cognitive psychology has shown us that individuals are more likely to use rules of thumb or heuristics when making social judgments (Kahneman, Slovic, & Tversky, 1982). In these contexts there is a tendency to overutilize information that is highly available (i.e., very salient or easily recalled, prototypic instances), or to overutilize default relationships between environmental cues that are highly similar or representative of some expected association (Tversky & Kahneman, 1974). By the same token the social judge is seen as underutilizing the various formal, statistical relationships between the individual case and the population as a whole (e.g., base rates; regression to the mean, etc.). Such normative information can often yield more appropriate judgments, such as assessing the dangerousness of a discharged patient when the base rates indicate that such individuals are unlikely to become violent (cf. Nisbett & Ross, 1980). Interestingly, the behavioral clinician seems just as likely to ignore such base rate information. In the classic Langer and Abelson (1974) study, behavioral clinicians, although less influenced by the priming label *patient* versus *job applicant* in their judgments of adjustment, were less likely to utilize the base rate

data. It has been suggested that in the absence of other information, an individual who is a patient could be more likely to have behavior problems (Davis, 1979). On the other hand, perhaps behaviorists are revealing a positive tendency to wait for such information before making a judgment that would harm the patient.

There is ample evidence that (nonclinician) judges rely on hypothesis- or schema-confirming strategies when making personality assessments. Snyder (1981) has shown through a series of studies that once an initial judgment is made about a person (in regard to a personality description or vignette), this judgment is used later even in the face of disconfirmatory evidence; that confidence in personality judgments increases over time, even in the face of disconfirmatory evidence. It is to be hoped that the behavioral model of personality, with its emphasis on situational interactions and response relationships, voids such tendencies. Overconfidence in biased judgments has been shown to correlate positively with the amount of available information but to be uncorrelated with the accuracy of medical (diagnostic) judgments (Elstein *et al.*, 1978; Oskamp, 1965). Einhorn and Hogarth (1978) have shown that such overconfidence in potentially biased and inaccurate judgments is due to the fact that when social judges make an estimate of judgmental confidence they base the estimates solely on confirmatory evidence. These cognitive strategies have been shown to create an illusion of validity in an individual's judgments.

Watts (1980) uses a metaphor from signal detection methodology, which may assist the reader in making some sense out of these biased judgment data, clinically. He suggests that experienced clinicians may be more likely to utilize "high risk," inferential strategies. They may be more motivated to lower their perceptual criteria (Beta, or response bias) and to utilize the presence of a relationship between perceived clinical data and a hypothesized schema that increases their clinical sensitivity (discriminatory ability or d'). However, such an increased ability to venture at best guesses is concomitant with an increase in the rate of false alarms, that is, the overall soundness of high-risk judgment may be compromised. If we borrow from Snyder (1980), Einhorn and Hogarth (1978), and Tversky and Kahneman (1974), we can come up with a conceptual integration of such ideas: if high-risk judgments are made that increase a clinician's ability to find "correct hits" (inferences) at the expense of false alarms or disconfirmations, and if disconfirming evidence is underutilized in subsequent judgments, then the possibility for an illusion of validity and overconfidence in illusory judgments comes full circle. In addition, if we add to all of this the very great possibility that clients may be unwilling to provide the corrective feedback necessary when an inference is made that does not feel right to them (due to demand, lack of knowledge, undue deference given to the clinician, unassertiveness, or maybe they are not given the chance due to leading questions, etc.), then the reinforcement of high-risk strategies becomes self-perpetuating.

The present analysis leaves more questions unresolved than answered. However, there would appear to be some value in incorporating models of decision making into behavioral assessment. Behavioral assessment includes measurement strategies and a vast set of assumptions regarding behavior, its organization, and the most valuable processes of discovery (hypothesis testing, for instance). How well clinicians utilize these rules (assessment integrity?) is a topic that would logically follow the

previous discussion, but for which there is little information. Whether following the rules makes any difference is a question only now being posed (Hayes, Nelson, & Jarrett, 1986).

INTEGRATION AND CONCLUSION

Behavioral assessment has developed in concert with behavior therapy mostly because of the latter's *sine qua non* being a dedication to objective outcome evaluation. This has resulted in particular concern for the fidelity of measurement. Practicing behavioral clinicians do need good outcome information, but they also need information allowing them to make the most appropriate treatment decisions. That is certainly not a novel observation; however, the idea that assessment must then be evaluated from this starting assumption has only just recently dawned. Clinicians require information that is correct and useful: that makes a difference to whether, what, and how they treat, and that influences their effectiveness. Some behavior therapy researchers seem to believe that therapeutic efficacy will be a property of the technique; we believe it will be a function of *treatment design by assessment information* interactions. Thus, we return to the symbiosis pointed out at the beginning of this chapter: assessment information provides the measure of treatment efficacy, but treatment efficacy provides the measure of appropriate assessment.

For these reasons we have tried to make clear the need for behavioral assessment to go beyond the properties of measurement devices to ask questions about *information*, a term that seems in keeping with the cognitive necessities of clinical practice. Sometimes the information needed might be a person's score on a test, but very often the information is more descriptive, such as the nature of the stimulus that elicits anxiety, or an accurate template of the environment for which the therapist is hoping to design more adaptive client skills. In this formulation we see considerable integrative potential. Idiographic and nomothetic traditions would not be at cross-purposes, because descriptions of the general (e.g., a finding from psychopathology research) becomes a very likely hypothesis that could be tested for its applicability to the unique circumstances of the individual client. Testing such individual hypotheses may not require quite as much rigor as testing general laws, although obviously some sort of protection is needed to safeguard the client from clinical hunches that could have serious deleterious consequences. This in turn provides some integration with professional ethics, the design of socially acceptable treatments, and clinical judgment.

There is another way in which the approach outlined provides integration. As we explained, the earlier interest in clinical judgment and decision making was rendered irrelevant because the model of the clinician's task used was an artificial derivative of psychologists' major assessment function being perceived as the assignment of diagnostic labels. Although that is a fairly complex judgment it can be task analyzed quite easily: it requires accumulation of the information that is most closely related to the diagnostic criteria and the application of the inclusionary and exclusionary decision rules of the diagnostic system. Once we abandon that paradigm, however, the alternative model for the decisions required is not at all clear. As we

explained, even simple prioritization of already identified target behaviors is an inappropriate assessment activity if clinical behaviors interrelate. Various scholars in behavioral assessment are struggling with ways of modeling or depicting the essence of the tasks clinicians must engage in. Kanfer, for instance, first identified the *categories* of information that might be useful (Kanfer & Saslow, 1969), and now is trying to articulate the nature of the processes for using that information (Kanfer, 1985).

With similar intent, Evans (1985) suggested that a systems perspective was helpful, in that the needed cognitive strategies for behavioral assessment involved constructing hypothetical conceptual models that place the internal and external influences on the client into some kind of logical, causal order. One clear advantage of this approach, as with systems models in general, is that any identified element in the system can be looked at in more fine-grained detail without posing any contradictions to the overall model. For instance, the systems conceptualization of a client might suggest that fear of open spaces is an element that causally relates to social behavior (not going out with friends), which results in restrictions of social contacts, that in turn make the client lonely and depressed. If the therapist then wanted to conduct a detailed analysis of the fear component, a more molecular model could be developed of its specific subelements, such as the precise stimuli in the environment that elicit fear (which Wolpe, 1986, correctly insists must be a primary assessment activity), the labeling repertoire imposed, subjective interpretation of the psychophysiological changes, and so on. Thus the systems model allows for a number of levels of influence and behavioral regulation to be considered simultaneously by the clinician. As of yet, however, the rules for constructing and verifying such complex individual models have not been worked out.

There is one additional way that a conceptual analysis of behavioral assessment as decision making provides integration, and that is in the wide array of different types of measure that has been brought within one assessment rubric. That rubric is, roughly speaking, the natural science perspective of behaviorism, although it is fashionable to argue that behaviorism no longer covers a set of agreed upon metatheoretical assumptions. We, however, are interpreting the term quite loosely, certainly to include cognitive constructs and personality constructs (Staats, 1986), but not including mentalistic or intrapsychical models of causation. An exciting variety of measurement strategies are successfully included within this behavioral perspective, resulting in a rich menu of measurement techniques and strategies. A clinical disorder framework discussing the psychophysiological assessment of anxiety and depression would rarely discuss simultaneously the naturalistic observation of mother-child interactions, the rating of skill deficits in persons with mental retardation, or the self-reported measurement of pain. Yet all these topics fit within behavioral assessment and have common measurement principles. That also means there is still a great deal of heterogeneity and no simple set of measurement strategies for clinicians to follow. Like behavior therapy itself, behavioral assessment cannot be conducted according to defined procedural rules—and will never produce psychological reports! In this chapter, then, rather than trying to specify how to do assessment, we have attempted to summarize some of the concepts surrounding the acquisition and use of veridical clinical information that seem to us most promising for effective clinical practice.

REFERENCES

- Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, *93*, 203–231.
- Barlow, D. H., Hayes, S. C., & Nelson, R. O. (1984). *The scientist-practitioner: Research and accountability in clinical and educational settings*. New York: Pergamon Press.
- Barrios, B., & Hartman, D. P. (1986). The contributions of traditional assessment: Concepts, issues, and methodologies. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 81–110). New York: Guilford Press.
- Bartlett, F. C. (1932). *Remembering*. Cambridge: Cambridge University Press.
- Bem, D. J., & Funder, D. C. (1978). Predicting more of the people more of the time. Assessing the personality of situations. *Psychological Review*, *85*, 485–501.
- Brown, L., Nietupski, J., & Hamre-Nietupski, S. (1976). Criterion of ultimate functioning. In M. Thomas (Ed.), *Hey, don't forget about me* (pp. 2–15). Reston, VA: Council for Exceptional Children.
- Burns, G. L. (1980). Indirect measurement and behavioral assessment: A use for social behaviorism psychometrics. *Behavioral Assessment*, *2*, 197–206.
- Cantor, N., & Kihlstrom, J. F. (1982). Cognitive and social processes in personality. In G. T. Wilson & C. M. Franks (Eds.), *Contemporary behavioral therapy: Conceptual and empirical foundations* (pp. 142–201). New York: Guilford Press.
- Cone, J. D. (1977). The relevance of reliability and validity for behavioral assessment. *Behavior Therapy*, *8*, 411–426.
- Cone, J. D. (1979). Confounded comparisons and triple response mode assessment research. *Behavioral Assessment*, *1*, 85–95.
- Cone, J. D., & Hoier, T. S. (1986). Assessing children: The radical behavioral perspective. In R. J. Prinz (Ed.), *Advances in behavioral assessment of children and families* (Vol. 2, pp. 1–28). Greenwich, CT: JAI Press.
- Corrigan, S. A. (1985). *A comparative study of behavioral treatments of bulimia nervosa*. Unpublished doctoral dissertation, State University of New York at Binghamton.
- Crowell, D. H., Evans, I. M., & O'Donnell, C. R. (Eds.). (1987). *Childhood aggression and violence: Sources of influence, prevention, and control*. New York: Plenum Press.
- Cutrona, C. E., Russell, D., & Jones, R. D. (1985). Cross-situational consistency in causal attributions: Does attributional style exist? *Journal of Personality and Social Psychology*, *47*, 1043–1058.
- Davis, D. A. (1979). What's in a name? A Bayesian rethinking of attributional biases in clinical judgment. *Journal of Consulting and Clinical Psychology*, *47*, 1109–1114.
- Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review*, *85*, 395–416.
- Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978). *Medical problem solving: An analysis of clinical reasoning*. Cambridge, MA: Harvard University Press.
- Epstein, S. (1979). The stability of behavior: I. On predicting most of the people much of the time. *Journal of Personality and Social Psychology*, *37*, 1097–1126.
- Epstein, S., & O'Brien, E. J. (1985). The person–situation debate in historical and current perspective. *Psychological Bulletin*, *98*, 513–537.
- Evans, I. M. (1986). Response structure and the triple-response-mode concept. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 131–155). New York: Guilford Press.
- Evans, I. M., Brown, F. A., Weed, K. A., Spry, K. M., & Owen, V. E. (1987). The assessment of functional competencies: A behavioral approach to the evaluation of programs for children with disabilities. In R. J. Prinz (Ed.), *Advances in behavioral assessment of children and families* (Vol. 3, pp. 93–121). Greenwich, CT: JAI Press.
- Evans, I. M., Meyer, L. H., Kurkjian, J. A., & Kishi, G. S. (in press). An evaluation of behavioral interrelationships in child behavior therapy. In J. C. Witt, S. N. Elliott, & F. N. Gresham (Eds.), *Handbook of behavior therapy in education*. New York: Plenum Press.
- Evans, I. M., & Nelson, R. O. (1974). A curriculum for the teaching of behavior assessment. *American Psychologist*, *29*, 598–606.
- Evans, I. M., & Wilson, F. E. (1983). Behavioral assessment as decision-making: A theoretical analysis. In M. Rosenbaum, C. M. Franks, & Y. Jaffe (Eds.), *Perspectives on behavior therapy in the Eighties* (pp. 35–53). New York: Springer.

- Evans, I. M. (1985). Building systems models as a strategy for target behavior selection in clinical assessment. *Behavioral Assessment*, 7, 21–32.
- Eysenck, H. J. (1982). Neobehavioristic (S-R) theory. In G. T. Wilson & C. M. Franks (Eds.), *Contemporary behavior therapy: Conceptual and empirical foundations* (pp. 205–270). New York: Guilford Press.
- Eysenck, H. J., & Rachman, S. (1965). *The causes and cures of neurosis*. London: Routledge & Kegan Paul.
- Farkas, G. M., Evans, I. M., Sine, L. F., Eifert, G., Wittlieb, E., & Vogelmann-Sine, S. (1979). Reliability and validity of mercury-in-rubber strain gauge measure of penile circumference. *Behavior Therapy*, 10, 555–561.
- Felton, J. L., & Nelson, R. O. (1984). Inter-assessor agreement on hypothesized controlling variables and treatment proposals. *Behavioral Assessment*, 6, 199–208.
- Glaser, R. (1981). The future of testing: A research agenda for cognitive psychology and psychometrics. *American Psychologist*, 36, 923–936.
- Goldfried, M. R., & Kent, R. N. (1972). Traditional versus behavioral assessment: A comparison of methodological and theoretical assumptions. *Psychological Bulletin*, 77, 409–420.
- Gottman, J. M. (1979). *Marital interaction: Experimental investigation*. New York: Academic Press.
- Gottman, J. M. (1983). How children become friends. *Monographs of the Society for Research in Child Development*, 48 (3, Serial No. 201).
- Hastie, R. (1981). Schematic principles in human memory. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario symposium on personality and social psychology* (pp. 63–84). Hillsdale, NJ: Erlbaum.
- Hay, W. M., Hay, L. R., Angle, H. V., & Nelson, R. O. (1979). The reliability of problem identification in the behavioral interview. *Behavioral Assessment*, 1, 107–118.
- Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1986). Evaluating the quality of behavioral assessment. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 463–503). New York: Guilford Press.
- Herbert, M. (1981). *Behavioural treatment of problem children: A practice manual*. London: Academic Press.
- Hollon, S. D., & Kriss, M. R. (1984). Cognitive factors in clinical research and practice. *Clinical Psychology Review*, 4, 35–76.
- House, A. E. (1980). Detecting bias in observational data. *Behavioral Assessment*, 2, 29–31.
- Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, 15, 336–352.
- Kaloupek, D. G., & Levis, D. J. (1983). Issues in the assessment of fear: Response concordance and prediction of avoidance behavior. *Journal of Behavioral Assessment*, 5, 239–260.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1982). Intuitive prediction: Biases and corrective procedures. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 95–128). Cambridge: Cambridge University Press.
- Kanfer, F. H. (1985). Target selection for clinical change programs. *Behavioral Assessment*, 7, 7–20.
- Kanfer, F. H., & Saslow, G. (1969). Behavioral diagnosis. In C. M. Franks (Ed.), *Behavior therapy: Appraisal and status* (pp. 417–444). New York: McGraw-Hill.
- Kuppermink, M., & Heimberg, R. (1983). Consequences probability and utility as factors in the decision to behave assertively. *Behavior Therapy*, 14, 637–646.
- Langer, E. J., & Abelson, R. P. (1974). A patient by any other name . . . Clinician group difference in labeling bias. *Journal of Consulting and Clinical Psychology*, 42, 4–9.
- McReynolds, P. (1979). The case for interactional assessment. *Behavioral Assessment*, 1, 237–247.
- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and review of the evidence*. Minneapolis, MN: University of Minnesota Press.
- Meehl, P. E. (1960). The cognitive activity of the clinician. *American Psychologist*, 15, 19–27.
- Merluzzi, T. V., Glass, C. R., & Genest, M. (Eds.) (1981). *Cognitive assessment*. New York: Guilford Press.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, 80, 252–283.
- Moos, R. H. (1973). Conceptualizations of human environments. *American Psychologist*, 28, 652–665.
- Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

- Nelson, R. O. (1977). Assessment and therapeutic functions of self-monitoring. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification* (Vol. 5, pp. 263–308). New York: Academic Press.
- Nelson, R. O., & Hayes, S. C. (Eds.). (1986). *Conceptual foundations of behavioral assessment*. New York: Guilford Press.
- O'Farrell, T. J., Cutter, H. S. G., Bayog, R. D., Dentch, G., & Fortgang, J. (1984). Correspondence between one-year retrospective reports of pretreatment drinking by alcoholics and their wives. *Behavioral Assessment, 6*, 263–274.
- Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Consulting Psychology, 29*, 261–265.
- Rugh, J. D., Gable, R. S., & Lemke, R. R. (1986). Instrumentation for behavioral assessment. In A. R. Ciminero, K. S. Calhoun, & H. E. Adams (Eds.), *Handbook of behavioral assessment* (2nd ed, pp. 79–108). New York: Wiley.
- Snyder, M. (1981). Seek and ye shall find: Testing hypotheses about other people. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario symposium on personality and social cognition* (pp. 85–116). Hillsdale, NJ: Erlbaum.
- Sobell, L. C., & Sobell, M. B. (1986). Can we do without alcohol abuser's self-reports? *The Behavior Therapist, 7*, 141–146.
- Staats, A. W. (1986). Behaviorism with a personality: The paradigmatic behavioral assessment approach. In R. O. Nelson & S. C. Hayes (Eds.), *Conceptual foundations of behavioral assessment* (pp. 242–296). New York: Guilford Press.
- Strosahl, K. D., & Linehan, M. M. (1986). Basic issues in behavioral assessment. In A. R. Ciminero, K. S. Calhoun, & H. E. Adams (Eds.), *Handbook of behavioral assessment* (2nd ed., pp. 12–46). New York: Wiley.
- Taylor, S. E. & Crocker, J. (1982). Schematic bases of social information processing. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario symposium on personality and social psychology* (pp. 176–204). Hillsdale, NJ: Erlbaum.
- Turk, D. C., & Salovey, P. (1985). Cognitive structures, cognitive processes, and cognitive-behavior modification: II. Judgments and inferences of the clinician. *Cognitive Therapy and Research, 9*, 19–33.
- Tversky, A., & Kahneman, D. (1974). Judgments under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.
- Voeltz, L. M., & Evans, I. M. (1982). The assessment of behavioral interrelationships in child behavior therapy. *Behavioral Assessment, 4*, 131–165.
- Voeltz, L. M., & Evans, I. M. (1983). Educational validity: Procedures to evaluate outcomes of programs for severely handicapped learners. *Journal of the Association for the Severely Handicapped, 8*, 3–15.
- Wampold, B. E., & Jenson, W. J. (1986). Clinical significance revisited. *Behavior Therapy, 17*, 302–305.
- Watts, F. N. (1980). Clinical judgment and clinical training. *British Journal of Medical Psychology, 53*, 95–108.
- Wilson, F. E., & Evans, I. M. (1983). The reliability of target-behavior selection in behavioral assessment. *Behavioral Assessment, 5*, 15–32.
- Wolpe, J. (1986). Individualization: The categorical imperative of behavior therapy practice. *Journal of Behavior Therapy and Experimental Psychiatry, 17*, 145–153.
- Yarrow, M. R., & Waxler, C. Z. (1979). Observing interaction: A confrontation with methodology. In R. B. Cairns (Ed.), *The analysis of social interacting: Methods, issues and illustrations* (pp. 3–28). New York: Erlbaum.

CHAPTER 17

A Response Process Model of Behavior

Kieron P. O'Connor

INTRODUCTION

There are different schools of thought on the current uneasy relations between theory and practice in clinical psychology. Theoreticians such as Eysenck (1981) take a procedural point of view and argue that clinicians are not rigorous enough in applying learning theory procedures. Clinicians (e.g., Cullen, 1983) argue that the problems are substantive, that everyday behavior unlike laboratory behavior is a complex business requiring multi- or even megavariate control.

Perhaps such a rift reflects a basic definitional question: Does behaviorism offer an adequate definition of everyday behavior? Is there any place for everyday behavior in current behaviorist thinking? Initially the question appears naive. The existence of behavior is after all the rationale for behavior therapy. Behaviorists must be able to define behavior or they could not identify types of behavior, analyze units of behavior, or change aspects of behavior—all of which clinical and experimental psychologists do regularly. However, the argument here is that behavioral therapy offers no consistently behavioral definition of behavior. Consequently the “behavior” of the experimentalist is not the “behavior” of the clinician, and the lack of definitional consensus embarrasses the theory in practice.

The initial sections review briefly the theoretical attempts to define behavior for clinical purposes and their relative failure. It is argued that in the absence of any consistent behavioral criterion for defining behavior, clinical psychologists have largely technicalized commonsense notions of behavior. The later sections propose that behavior can only be defined behavioristically if a response process model of behavior is adopted. Interestingly, theorists from divergent positions seem to be converging on such a model. The response process model views behavior as an emergent property of background response activity; it adopts a contextual approach to defining behavior and suggests that behavior is best understood by describing the context of processes that made its emergence possible, rather than by attempting to explain it as a discrete

response to a source event. Everyday behavior is then classified according to equivalence of response processes rather than defined in relation to fixed sensory coordinates.

THE PROBLEM OF BEHAVIOR

The lay reader of psychology journals might even be forgiven for thinking that people do not seem to do things anymore; they behave at them. Talking, walking, jumping up and down are respectively verbal behavior, ambulatory behavior, out of seat behavior. The purpose of classifying activity as behavior is to locate the activity within the behavioral domain, as a necessary prelude to applying behavioral methods of interpretation and intervention. Talking is regulated by physio-chemico-socio-linguistic factors, whereas verbal behavior is behaviorally regulated. Jumping up and running about likewise may have multiple determinants, but "out of seat" behavior is under behavioral control.

Labeling an activity or a set of activities as behavior has implications for behavioral management, but it does not furnish us with any additional parameters about the activity nor specify it further. It does not for example allow us to replace the commonsense categorization of activity with a more exact metric. Someone is talking and therefore they are verbally behaving. They are not talking *or* behaving. The term *behavior* then appeals to and depends on prior lay categorization of activity. Someone who is reaching out to pick up a glass might be said to be behaving, but by knowing that they are behaving we do not understand what they are doing other than reaching, or that they are reaching in a particular way.

COMMONSENSE DEFINITIONS OF EVERYDAY BEHAVIOR

The concept of behavior is an organizing construct that places boundaries on a person's actions. The question is according to what logic are the boundaries of behavior commonly defined?

The pioneering aim of Eysenck (1980) and others in replacing psychotherapy by behavior therapy was to give therapy an empirical base. Identifying a response within an experimental set-up is indeed an empirical business. The person's behavior is represented by his or her response deduced from observation of the stimulus. In everyday life the deductive logic is reversed. Behavior presents itself. The behavior is assumed to represent a response and the stimulus or reinforcement is inferred from observation of the behavior.

Most psychologists would argue that in identifying behaviors they are acting empirically because they are identifying responses. But without a predefined stimulus-response (S-R) contingency there can deductively speaking be no response. Because everyday behavior is always identified before knowledge of any associated contingency the behavior cannot have been derived empirically but by some other method; and because behavior always coincides exactly with lay notions of activity, the method must be commonsense inference. So when the clinical psychologist deals with everyday behavior, she or he is confronted with behavior in a culturally, not a behavioristically,

defined form. But the cultural designation of behavior depends on prior ascriptions of wilfulness and personal accountability of a person that are moral not scientific. Behavior as a lay construct then is often a judgment on, not a description of, activity. So saying, for example, that a person behaved appropriately or inappropriately will often be a sufficient classification of the person's activity without further specification.

The contextual nature of such lay classification of behavior means that different activities may be classified as identical behavior on separate occasions. When considering the problem behavior of a child, for example, the criterion for the inclusion of activity in the class or unit of behavior to be analyzed will focus on what is seen to be disruptive. There may be several distinct activities involved in the problem behavior but these will be lumped within the same class of behavior (Eyberg & Robinson, 1985; McFall, 1979).

Conversely the child told to "behave yourself" will be expected to exert a control and regulation over physiological activity not on the basis of any specialized skills or knowledge the child may have of his or her activity, but because self-control is implied by the moral judgment that the child's activity is self-willed behavior. In general the nonconforming person's actions are attributed to individual rather than collective responsibility. The problems of a child labeled "behaviorally disordered" are more likely to be approached behaviorally than say, medically, as compared to the problems of an unlabeled child (Fernald & Gettys, 1980).

These limitations work against the behaviorist's practical aim of being as specific as possible about process. *Process* is defined here as the method of operation, the series of actions of a given behavioral act. The behaviorist might also observe a clash between applying behavioral analysis to novel areas, such as physiological activity or medical complaints, and maintaining a definition of behavior within normative bounds. These bounds denote, for example, a consensually rather than scientifically understood degree of wilfulness that makes it socially inappropriate to talk of "cancer behavior" or "heart attack behavior" without transgressing an accepted division of action into self-willed and other-willed, and ascribing an inappropriate moral culpability. Hence it is essential that the behaviorist redefine behavior, behavioristically, to allow the behavioral domain to be built up from activity classified on an empirical basis.

The problem is that the early theorists never considered complex behavior, but self-evident responses and reflexes, the sources of which were readily available. Watson and Skinner, for example, were less concerned to define behavior than to define the behaviorist rationale.

BEHAVIOR AS REPRESENTATION

Classical theories of behavior view behavior as a representation or manifestation of some other more profound event or process. Their approach to defining behavior is to explain it away as the product of something else. Conviction about the sensory basis of all behavior had led to a lack of concern with behavior itself. The emphasis has been placed on defining the mechanics of the assumed source of behavior rather

The *constitutional* approach generally attempts an intraorganismic explanation of behavioral events, and ranges from the reductionist school to the typologists' approach. The naive reductionism of both the physicalist and mentalist kind that behavior is "nothing but" a collection of neurons firing, or that behavior "comes down" to a mental attitude is widely recognized as misconceived. It confuses levels of analysis with levels of explanation (Markham, 1979). But some constitutional theorists (e.g., Gray, 1970, and Beck, 1976, for example) still tend toward hierarchical explanations. Precedence is given to one area of function and behavior is seen as the end point of a series of stages that make it up.

Conversely, typological theorists like Teplov, Eysenck, and Strelau view behavior as a property or ascriptive of a certain type of organic structure. This puts behavior into a separate logical category from, say, the physiology of the person. A behavior becomes a predicate rather than a product of the nervous system.

The third approach is the contextual approach advocated here, which views behavior not as a function, a representation, a correlate or a sign of anything but as a symbol complete in itself and expressing a logical order within a person's actions.

SENSORY MODELS OF BEHAVIOR

All sensory models take for granted an ordered relationship between stimulus and response (Fig. 2). Any knowledge gained from the environment is preceded by a stimulus property transmitted to and received by the person. Behavior then is entirely a function of the stimulus plus or minus transfer function error.

The main problem with the physicalist sensory theories is that much more power is ascribed to physical events than can be explained physically. Though the stages of the sensory processing are physically marked out in accordance with transducer models, the actual processes involved remain metaphysical and glossed over

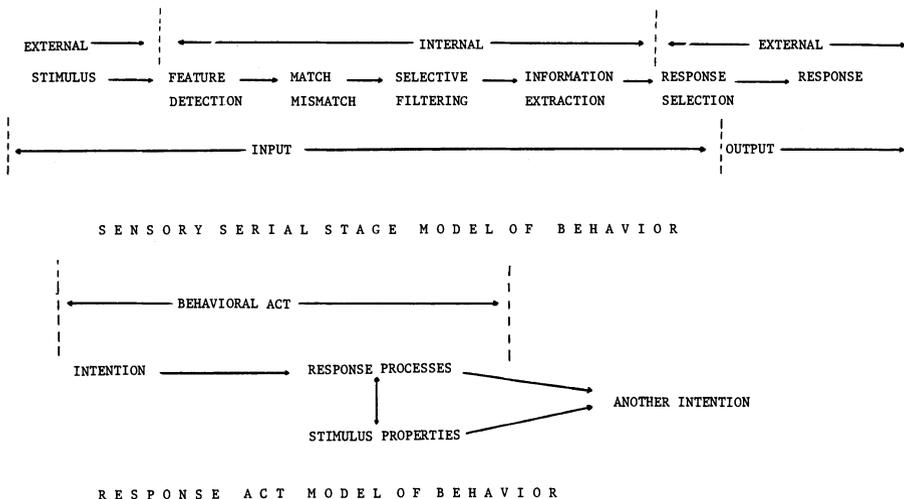


FIGURE 2. Schema of the sensory model of behavior and the response act model.

with physicalist metaphors. Response events are said to be “stimulus driven” (Treisman & Gelade; 1980). (What is driven by whom?) Stimuli are “mapped onto” responses (Fisher, 1982). (Who and where is the cartographer?) Investigators are content to accept the fact of transmission, registration, and storage, and leave the how to future research.

The concept of a fixed sensory threshold to physical properties assumed by physicalist sensory models was challenged by signal detection theory (SDT). The separation of perception by SDT into separate discriminability and response bias components better accounts for variations in detection. In recent developments Green and Birdsall (1978) suggested that these components may not be distinct and that every stage of perception may involve decision making. The concept of an all-or-nothing stimulus threshold is replaced with a continuously variable and active construction of detection.

SDT highlighted the inadequacy of passive sensory models to explain motivation to respond or acquisition of knowledge without active psychological involvement of the person. This insufficiency was inevitable because in the event the phenomenon to be explained was not the physical characteristics of the stimulus, but the characteristics perceived by the person. The nonphysical basis of stimulus selectivity is generally built into a sensory stage or channel that precedes awareness. But the stage remains hypothetical. As Laming (1985) puts it, “There *must* be a preconscious, analytic stage not available to introspection when those boundaries and other features of the stimulus field are identified. This prior stage is the domain of sensory analysis,” (p. 462). But then to what stimulus does this prior stage respond, or do we have to have a pre-preconscious stage? If this stage is part of the response then clearly the response is not sufficiently explained either by the stimulus or by any physical properties of the response system.

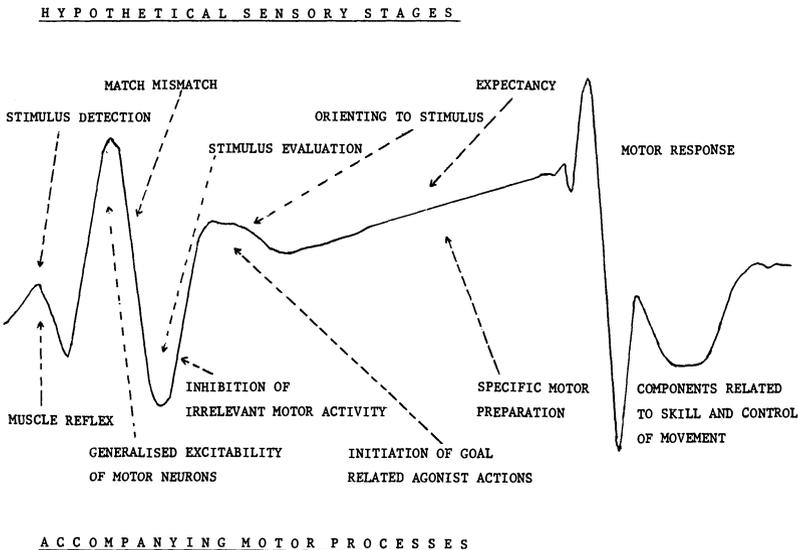


FIGURE 3. Alternative sensory and motor views of cortical event related potentials.

The attempt by naive sensory models to mate psychology and physics has saddled it with a permanent demarcation problem. At what point is information transmitted from outside to inside? When does a sensory event transform into a sensory property and then into a response event?

The closer one looks for an example of a purely sensory function the more precisely one discovers the active participation of the subject in its genesis, even among micropsychophysiological events (see Figure 3). For example, the peaks and troughs of the evoked response are seen to support the classic sensory picture of physical detection stage, match/mismatch stage, registration stage, semantic evaluation stage, etc. But early components now appear to be as cognitive as later components (Parasuraman & Beatty, 1980). Furthermore, component differences change with practice and skill (Sabat, 1979). Research on orienting has highlighted the difficulty of separating components according to sensory function, and has shown that quite abstract properties can elicit large initial responses in the subject (Bernstein, 1979; Siddle, 1985).

The passive sensory model of behavior cannot then provide an adequate account of the processes involved in responding to sensory events and has focused on response parameters irrelevant to response processing.

The serial processing model also assumes that movement is modified by a process of change in single elements of the act. But in general motor actions start from an abstract level and proceed through decreasing degrees of freedom to a specific concrete action (Turvey, Shaw, & Mace, 1978). Control of a given movement represents the capability to change rapidly from an abstract to a precise joint/muscle combination within a goal-directed response (Newell, 1978). This is in direct conflict with sensory approaches, which view information processing as starting from a discrete stimulus level and progressing to a more complex format.

Because the sensory model does not adequately account for the processes of observed behavior, it cannot offer an empirically based definition of behavior.

CONDITIONING MODELS OF BEHAVIOR

As generally treated within conditioning models, behavior is a response whose function lies in its adaptive value to the organism. All behavior is thence an adaptation or maladaptation to the environment. Conversely, properties of the stimulus environment are only responded to insofar as they affect the evolving adaptation of the organism. In both classical and operant conditioning it is the *implications* of environmental contingencies that principally determine their associative strength.

Identifying behavior as the effect of stimuli or reinforcement requires that stimuli, in the case of classical, and reinforcement, in the case of operant conditioning, must be defined independently of the response to avoid the trivializing consequences of circularity.

But the various conditioning paradigms permit stimuli to precede, follow, or coincide with behavior, and to be internal or external states, visible or invisible (Grant, 1964). Theoretically, then, contingencies need not be related in any absolute way to the occurrence of behavior. Additionally, the response itself may be both the reinforcer

and the stimulus for its own occurrence through proprioceptive refference or secondary reinforcement (Perkins, 1947).

On the other hand completely independent categories of response may be conditioned in addition to the targeted response, including irrelevant responses. The classic example is Albert's thumb-sucking acquired during fear acquisition (Watson & Raynor, 1920).

Grant (1964) states that no S-R connections exist in isolation. Each is embedded in a complex matrix of behavior and when we isolate a particular S-R connection for logical consideration we can never really isolate it from its matrix. This idea of a complex continuum of functional stimulus impact in place of an all-or-nothing response has become increasingly popular in conditioning explanations (e.g., Bor-kovec, 1979).

In practical terms this means that identifying behavior as a conditioned response can offer no firm predictions on the temporal or spatial limits of the response or its generality. It is hence very difficult to say exactly where it begins or ends and also to separate antecedents and consequences from concomitants of the response. Indeed structural analyses of conditioned response processes, as in Martin and Levey's (1969) work on response topography for example, have shown that in large measure response processes develop independently of fixed contingency constraints and take on a life of their own. This means either that other factors within the contingencies are responsible for the response and/or that there may be more to the response itself than is permitted by the S-R contingency.

The fact that verbal stimuli, instructions, and general background features of the laboratory situation can vastly facilitate acquisition of a response (Prokasy, 1965) may suggest not that there are other cognitive features within the conditioning paradigm, but rather that the conditioned response is one aspect of a larger behavioral unit that includes compliance in entering the conditioning laboratory in the first place. In other words the passive conditioned response is part of a wider behavior pattern on the part of the subject in agreeing to participate in the experiment, and accepting the rules of the experimenter.

Hence we are replacing the conditioning model of a discrete response elicited by association with a contingency with an active skill model that views the response emerging against a background behavioral complicity on the part of the subject that allows it to emerge. If the rules for this complicity change so does the response, and in a more dramatic fashion than by altering the S-R contingency (Jung, 1982).

None of this is to detract from the important empirical work that has emerged from conditioning laboratories, but this may be said generally to demonstrate conditioning as one and possibly the first example of the conditional nature of behavior, rather than the ultimate conditioned nature of all behavior. Conditioning theory fails, however, to offer a consistent logic by which to define everyday behavior in conditioned response terms. Understanding the purpose of behavior in conditioning terms and in particular classifying behavior as functional or dysfunctional according to adaptational criterion must be challenged.

Eysenck (1981), for example, following Mowrer, defines the "neurotic paradox" as a "self-perpetuating and self-defeating" behavior. This is so only if dysfunctional

outcome is defined according to adaptive values prescribed by an experimenter-defined contingency. If behavior serves another purpose undefined by the contingency then a nonrewarding outcome becomes more difficult to isolate.

In practice, when dealing with the acquisition of new behavior, therapists tend to adopt a skills model of behavioral training that stresses competence rather than adaptation. As McFall (1979) points out, most conditioning models assume that the response is already at least potentially in the person's repertoire and only needs to be facilitated or inhibited. In the skills model, however, clients are seen as requiring new skills to achieve valued personal goals in specific situational tasks. There is nothing in the skills approach that prevents neurotic behavior being viewed as a competent skill, serving a positive purpose in some situations.

CONSTITUTIONAL APPROACHES TO BEHAVIOR

PHYSIOLOGICAL THEORIES

The constitutional approach has recognized the self-determination of much of behavior and viewed it as a product of intraorganismic processes. Many of the criticisms of the sensory model remain valid for reductionist physiological theories of the constitutional mould, because a physiological state represents little more than an internalized stimulus, preceding and determining the behavior. As in the sensory model the source state is inferred from the response, and is even synonymous with it. A generic example of this type of constitutional hypothesis is the classic arousal concept. This was introduced by Duffy (1957) as an attempt to unify the fragmented physiological concepts of drive, energy, excitation. She postulated arousal as a single unitary dimension, having a linear relation to physiological indexes of activation and to performance.

The predictive value of arousal as a constitutional explanation of behavior has become lost in a quagmire of nonspecific neurophysiological operations. Gale (1981) notes that arousal is often confused with a source of stimulation, an endogenous variation, an experienced drowsiness or alertness, a correlate and consequence of action, a drive, and motivation—all of which clearly involve different hypothetical behavioral processes.

The *typological* approach has generally eschewed sensory causation for an integrated systems model of individual behavior. Eysenck's (1967) use of the term *arousal* in his revised typology theory is very much as a self-regulatory factor, and in contrast to its general use as a stimulus source state. The systems approach of the typologists accepts a more flexible reciprocal relation between sensory and motor function. The active role of behavioral process in modifying sensory events is implied in Eysenck's and Teplov's theories and is essential in their predictive accounts of individual behavior and biological regulation.

Recently, Strelau, in attempting to be more specific about process and to reconcile Pavlovian and Eysenckian conceptions, has proposed style of action as the key expression of temperament. According to Strelau (1983), *style of action*, which develops

under environmental influences, is considered to be one of the regulators of stimulation. Style of action is a goal directed activity, made up of basic and auxiliary activities. Basic activities directly modify the results of action. Auxiliary activities organize the conditions for the performance of the primary actions and modify conditions during performance, and the ratio of auxiliary to primary activity undertaken is an expression of the individual's type of reactivity.

This approach to explaining behavior is a motor skills approach, centered around the specific actions of the person as self-contained, self-determined phenomena. Even here though individual actions become typed along a general hypothetical dimension derived largely from classification of task factors rather than person centered intentional action. Consequently we are presented with a motor but not a behavioral definition of behavior.

COGNITIVE THEORIES

The cognitive argument developed as a reaction to the implausibility of the S-R response model of behavior, but failings of this response model were viewed as failings in stimulus definition rather than behavioral definition. Consequently the early cognitivists (e.g., Beck, 1976) accepted without question the S-R response limits placed on behavior. They sought only to supplement stimulus cause with causal mental variables equally as hypothetical as external variables. But because vague and ill-formulated intangibles, such as awareness or cognitive set, could not themselves be defined, they could not help define behavior. As Coyne (1982) pointed out, there is nothing that can be regarded as a prototypical conception of cognition. Where cognitive terms are operationalized in clinical practice they cannot be divorced from behavior. Someone who is thinking, or talking, is also behaving in an observable way and cognitions have been most manageably defined by considering them in terms of response propositions.

The work of Lang (1980) is an example where covert processes have been guided by specific response imaging in order to produce physiological effects. In general the clinicians among the cognitivists have sought to redefine behavior rather than stimuli; they have not revoked but extended the behavioral domain into new realms, and at the same time tended to become more specific and less hypothetical about response processes.

Meichenbaum (1978), for example, views introducing cognitive processes as involving a shift in emphasis away from discrete situation-specific responses to a concern with skills that can be applied across response modalities, situations, and problems. Training clients in self-instruction or imaginal behavior enhances generalization of coping responses. Here Meichenbaum is extending and specifying manageable behavior, not denying or explaining away its existence.

Further evidence of the "behaviorization" of cognitivists is the recent interest in adopting the Bayesian decision-making heuristic as a way of operationalizing subjective belief (Kahneman, Slovic, & Tversky, 1982). This heuristic is most definitely behaviorist (Savage, 1972). Savage emphasizes that the importance of a behavioristic outlook is its emphasis on consequences, that is, its goal-directed or forward purpose.

In these limited examples from the clinical-cognitive debate, behavior has not been explained away by cognitive variables but the narrow definition and organization of S-R defined behavior has been challenged. The focus here has shifted to the processes of behavior themselves that, independent of stimulus events, may be actively responsible for their own regulation. This is the position of the motor model and is based on evidence for the independent role of response processes in defining behavior. A resume of this evidence and its implications follows next.

BEHAVIOR AS RESPONSE CONTROLLED

Behaviorally it is quite evident that response (in terms of motor output) is present at all stages of processing. Work on spinal activity during foreperiod preparation preceding responses indicate a longer muscular pretuning than could be acceptable to sensory models (Loveless, 1979; Turvey *et al.*, 1978). Muscle activity has been seen not only as a precursor or accompaniment to sensory processing but also as a consistent (though uncontrolled) constituent.

The regulatory role of voluntary and involuntary movement is evidenced by the gating effect of somesthetic potentials (Jones & Hulme, 1976) and motor potentials (Papakostopolous, 1980) on subsequent input, suggesting that even minor motor movement may enhance or diminish sensory processes and so act as a screen to incoming events.

Investigation of motor output as a reflection of cognitive activity has been most notably pursued by McGuigan (1978). His thesis is that the complex muscle response patterns that accompany cognitive acts in speech and skeletal musculature are essential to generate neurophysiological codes for cortical processing, and he has demonstrated the presence of muscular activity during several apparently passive, silent procedures such as listening, reading, and thinking. He concludes that mental processes are sequences of bodily events that occur under specifiable external and internal conditions and nothing more. In principle, when we have measured all bodily events according to McGuigan, we have specified the mental process in question. Such evidence of course speaks against any temporal separation of input and output processes and suggests that motor activity may actually be part of information processing rather than just an addition to it. If indeed cognition and sensory events may be adequately represented by motor activity, as McGuigan suggests, it seems reasonable to doubt their temporal dissociation and to suggest that sensory events may be entirely generated by response activity.

Such a point of view has expressed itself in a resurgence of interest among psychologists in the motor model of mind, for example, Weimer (1977), whose main thesis is that motor action is an implicit requirement for stimulus evaluation. This view has been forcefully advocated by Sperry (1969), who suggested that mental activities develop out of and in reference to overt action and therefore analysis of output would tell us more about mental processes than analysis of input. But the motor view has run as an undercurrent to the sensory approach for approximately 100 years.

The motor model was originally formulated by Munsterberg (1989) as an efference theory of perception. Munsterberg held that the vividness of conscious experience is a direct function of joint and muscle sensation. The motor discharge is necessary before any central activity corresponding to perception or consciousness can take place. The spirit of Munsterberg's theory has remained encapsulated in the James-Lange theory of emotion, where muscular changes are seen to account for emotional differentiation; in Freeman's (1948) bioenergetic theory of the muscular control of nervous energy; and even in more recent attribution theories linking somatic change with emotional labeling (Schachter, 1971).

BEHAVIOR AS SELF-REGULATED ACTION

If behavior is best viewed as a form of motor skill, then the search for behavioral process should be directed to a closer study of the constitution of motor skills themselves. However, such study produces no evidence that motor action is enacted through the coding of particular muscles.

Several authors have noted that action invariance across different muscle types must indicate a centralized rather than peripheral representation of the image of action (Turvey, 1977). It is generally accepted that the preselection of movement, or of a central image of action is necessary for movement organization, and that preselected movements may be accurately reproduced whether or not peripheral cues for motion are present. In other words, relational coordination rather than discrete movement governs preselection mechanisms.

According to a relational motor model, behavior is defined as motor process, and the unit of analysis for all behavior is the self-regulated motor act. "Sensory" functions from the motor point of view must be considered as behavior and hence be organized relationally. Feelings, sensations, and perceptions are viewed as acts, in the same way as muscle movements are acts. Each act is considered not as a discrete, isolated coded response to any internal/external stimulus element, but as part of a relational structure that extends beyond particular acts to include all acts of the person at any one time. All aspects of the person's response serve one organizing aim or intention rather than a series of separate functions. Turvey (1977), for example, has proposed a viable relational model of sensory processes to account for the perception of movement. The perception of velocity and acceleration, according to Turvey, is based on transformational information detected over time and not on the discrimination of elementary aspects of spatial and temporal position.

BEHAVIOR AS A RELATIONAL ACT

The motor act model then offers a plausible account of any behavior as an act, but an act organized in relation to other acts at the time and constructed by centrally coordinated intention.

The problems for the motor act model of behavior are in deciding the form of the relational organization and how one derives relational definitions of behavior

from observation of behavior. The problems are highlighted in Liberman and Mattingly's (1985) recent motor theory of speech behavior. The aim of the speech-act theory is to go beyond the idea that speech perception requires an arbitrary association of signal with phonetic category and a correspondingly arbitrary progression from an auditory stage to a superseding phonetic label. Speech variation cannot be accounted for by mediation and ascribed to diverse stimulus sources or to discrete response elements. Rather, all elements of speech overlap to form a linguistically significant gesture. Communication is automatic through immediate innate perception of the gesture that exerts central invariant control over the articulation of both speaker and hearer.

But where do we find these invariant properties? Liberman and Mattingly are unclear. "They are seen not as peripheral movements but as more remote structures. These remote structures correspond to the speaker's intention" (p. 6). But how do we recognize this intention? "The intention is not conveyed in any one element of speech nor is it directly observable" (p. 12). But at least we must be able to define the gesture that contains the intention that corresponds to the invariant properties. Alas, no! "It is no simple matter to define specific gestures rigorously or to relate them to observable consequences. Yet invariant gestures there must be" (p. 20). The latter sentence encapsulates the authors' dilemma. They know exactly from a motor viewpoint what behavior cannot be, but they are unable to offer any further insight into specific processes apart from the general notion that self-regulation must take a goal-directed relational form. Hence we are presented with processes and properties derived from negative inference. Remote structures must control movement because local ones cannot. Information must be innate because it cannot be learned through mediation.

Liberman and Mattingly (1985) admit that they "do not understand the system that computes the topologically appropriate version of a gesture" (p. 23). Hence their response processes remain as hypothetical as the mediational model they seek to replace.

BEHAVIOR AS AN ECOLOGICAL ACT

The work of Gibson (1966, 1982) is based on the recognition that a functional approach to behavior must take account of the relational nature of behavioral acts. But like other relational models Gibson is unable to operationalize this relation in any positive way.

His main theme is that all behavior is a (motor) act and that because the act is organized relationally everything is in on the act, which accounts for the ecological immediacy of behavior. Gibson points out that the stumbling block of all what he terms establishment theories of behavior is the mediational stage assumed to exist between the world and the organism in order to convert external properties into internal codes. Such a stage, according to Gibson, loses the essential immediacy of behavior: its "intuneness" and its ecological validity. The establishment theories also inevitably fragment the person into separate response systems when in fact the person responds as a whole. We do not simply see with our eyes but with our eyes in a

certain position at a certain height and accompanied and moved by the quality of our body movements.

The "effective stimulus" for Gibson is the appearance of the environment, and the appearance becomes a property instantaneously of the organism and the environment. In Gibson's terms the environment affords certain possibilities to the person by virtue of the person's movement towards the environment. The notion of a discrete stimulus is replaced by a perceptible form. Properties of the environment are picked up to match the person's position simply because they are implied by a particular physical attitude of the person. Information is not given or transmitted to any particular sensory system; on the contrary it is ecologically present in the physical layout.

The effective stimulus is dynamic, in that it contains information on the person's action toward the environment, so that it is relational properties that are perceived. A stick is "climbupable" at the same time that a person climbs up it, a garden path is walkupable at the same time as the person walks up it. The limits of these dynamic second-order properties are controlled by structural invariants that designate physical limits of action.

Gibson's strength is to have outlined the concerns of a relational behaviorism and in particular the ecological relevance of a response. But he does not outline any plausible response processes. There is also a curious asymmetry in the holistic relation of environment and person. It cannot be predicted what information will be afforded by the environment until an individual has behaved toward it. A stick is climbupable only when someone climbs up it. Behavior is observed, and only then defined in Gibsonian terms. This is exactly the problem referred to in the introduction, namely the need to redefine behavior without having to rely on prior lay definition.

The contradiction of the Gibsonian approach is that it attempts to answer problems derived from realism with more realism. In fact Gibson's direct realism might be more correctly retitled *metarealism*, because it is an attempt to resolve the basic realistic dilemma of matching an absolute world with a relative response by shifting the match/mismatch to a higher order set of invariants. Under the guise of expanding the response to ecological parameters, Gibson in fact pushes the response further back into a hypothetical realm.

The aims and intentions of action according to Gibson are embedded in the concrete structure of the world and remain apart from and unseen in behavior because they are organized at a higher unobservable level. This is not an empirical approach to defining behavior and not surprisingly it gives us no information on response process.

Gibson's definition of behavior remains a functional one. Behavior serves a functionally adaptive purpose to the environment here and now. But because behavior and the stimulus environment are one there is no independent criterion of adaptation, so all behavior is adapting to itself and hence all behavior is adaptive. This clever circularity, however, leaves us with no predictability about actual behavior precisely because behavior is an independent process, not a correlate of the environment. In order to define behavior behavioristically the relationship between the person and the world must be centered on and defined by response processes. The relational organization that we have discussed as a plausible model for action must be described relative to response processes, not to external constraints. This can be achieved by

a relativist but not a realist approach to behavior. Gibson's ecological realism is very much in the American pragmatic tradition of James, Pierce, Dewey, and Mead, rather than the European rationalist tradition of Husserl, Merleau-Ponty, and Sartre. Pragmatic utilitarian realism considers only one reality and defines behavior as fixed in a time-bound context. As James (1970) put it "on the pragmatist side we have one edition of the universe; on the rationalist side a universe in many editions" (p. 113).

The rational relativist does not dispute the real but does not accept it as an absolute metric by which to define things. Rather, the concern is to reveal the manner in which the real is real. This means the here and now is defined in a wider context of possible worlds and possible behavior.

THE RESPONSE-CENTERED APPROACH TO BEHAVIOR

If behavior is placed in the context of possible response processes rather than a fixed stimulus world, response processes can be understood independently, without stepping beyond the boundaries presented by the behavior itself.

Previous attempts to define behavior tied behavior closely to hypothetical models of how the behavior is produced. A single behavioral act could not be viewed as significant by itself, but could only be significant when consistently interpreted in terms of some other unit, such as an environmental, physiological, or a temporal unit. The uniqueness of a single unrepeatable act of behavior is hence lost in attempts to establish it as a repeatable hypothetical effect. So in practice such approaches to behavior can never rival commonsense definitions of behavior, because their technology takes them increasingly away from the processes of the individual action that must be the unit of everyday behavior.

We can therefore outline some basic requirements for a method that seeks to define behavior in terms of behavioral processes alone. The basic unit or framework for viewing any behavior must be the actual observed behavior itself. In other words the boundaries of any action must be decided by measures of the action and not by sensory, stimulus, or time markers. Also, the method must respect the unique intentionality of any individual act.

Any action no matter how small must be seen as an expression of the person's overall position in the world. Thus all aspects of a person's behavior are related in a contextual rather than causal fashion: parts of the act cannot cause one another. Immediately we begin to explain a phenomenon in terms of cause and effect we step beyond the phenomenon's processes and become hypothetical. Instead we can view acts as emerging from their intentional context. This emergence is understood by *description*, not *explanation* of the context. Explanation in terms of cause and effect generally aims to ascribe a significance to something that it does not already possess. But in the relativist position any response action is already significant by virtue of its occurrence. We do not have to justify its presence or tamper with it further. The job of the contextual approach is to typify this significance by describing the context of response processes that made the acts' emergence possible.

So, in practice how do we meet these theoretical requirements?

First, the context we are seeking must be made up of behavioral actions, and second, it must be centered around the unique action under consideration.

We can meet these requirements by considering the response in the context of alternative possible responses that the person could have made and has made on other occasions. Any action can be considered as the center of an action space made up of a class of alternative responses. More specifically, the space might be considered as a likelihood or possibility space, where the maximum likelihood is necessarily ascribed to the actual response act that occurred. On each side of this maximum are other possible actions with decreasing likelihoods of occurrence. The class is established on the basis of response equivalence, not stimulus or temporal equivalence. Hence we are looking for acts that share a common behavioral dimension, that is, behavior that involves similar motor acts, thought acts, speech acts, etc. Such acts can then be scaled as precisely as possible by measuring how much of the common behavioral dimension they contain. Behavioral measures may include physiological measures, or solely observational measures, and of course there may be more than one dimension. Initially, I may construct a class of behaviors around a response act according to a fairly gross behavioral criterion, and then redefine the dimension as I find out more about which measurable activities typify precisely this response act in this class of possible actions.

For example, a certain head action involved in dizziness behavior may be initially classed as one of a class of general head movements, and then as one of a class of head movements with a specific angle of rotation as we find that degree of head rotation is the behavioral parameter most likely to typify dizziness.

The aim is to find a class of independent acts as closely equivalent to the action involved in a behavior, but that differ from it in intentionality. The behavioral dimensions common to the class are quantified according to actual parameters. If the quantitative information is not available from past acts performed by the person then the person may have to perform the equivalent actions in order to gain the relevant information, to create a suitable response class.

This can be illustrated by considering the behavior of reaching to pick up a glass. The intentionality of the reaching act is revealed by considering it in the context of all possible equivalent reaching acts of the person, past and present.

If I reach out to pick up a straight glass, this act of reaching is one of a class of reaching behavior. They are all unique and independent but nevertheless potentially equivalent in response terms because they share the process of reaching. Forming a distribution of response acts as equivalent as possible in process terms to my particular reaching act constructs a context that typifies the uniqueness of my act in process terms.

If I am interested in revealing the response process typical of the act of reaching out for a straight glass, then comparing this act with reaching out to pick up an umbrella or to pick up a book does not afford a very precise equivalence. Considering the acts of reaching for a round glass, a brandy glass, a wine glass, etc., affords greater precision, whereas considering acts that reach over the same distance gives even greater sharpness to my distribution of equivalent and possible response acts. I might seek further information about these acts in terms of muscle activity or of timing of flexor-extensor movements, and examine these measures across my response

class, in order to see if there are any unique parameters that typify my reaching for the straight glass.

In other words, by seeking what is common to my class of reaching acts I establish what is unique to any particular one. This will enable me to typify that act and ascribe a high likelihood to the occurrence of that act given the presence of certain empirical response processes. If I find that a certain strength of grip which has quantifiable values typifies my act of reaching for a straight glass and want to explore the significance of this component further, I might treat grip as a response act and look across response-equivalent classes involving heavy or light grip, and then look for parameters—environmental, physiological, or psychological—that typify this particular grip action.

The response processes reveal the specific way in which a person in a single act expresses an intentional relation within the world. But the logic of the approach requires that the context of response processes is built up empirically by discovering new information about the activity in order to define it more precisely.

RESPONSE PROCESS ANALYSIS AND CONVENTIONAL FUNCTIONAL ANALYSIS

There are major differences between functional analysis as conventionally practiced in clinical psychology and the present response process analysis. Conventional functional analysis is concerned to establish stimulus or situational equivalence. Environmental antecedents, and consequences of a behavior (ABC analysis) are charted over several occurrences of the behavior to unravel variables that show the determinants of the behavior and the function that the behavior serves in the environment. In practice, however, associated stimuli often turn out to be too general or too variable to plan intervention, and treatment may depend heavily on speculation about source, a point made by Owens and Ashcroft (1982).

In response process analysis, actions do not have to be functionally related to stimuli in order to plan specific interventions. On the contrary, because stimuli or perceptions of stimuli are themselves acts and hence related to overall intentionality, nothing is to be gained by according to them a special status. Indeed, completely separate stimulus perceptions may be considered response equivalent if they are part of the same intentional action.

Crucial to process analysis is the respect of individual actions: individual behavioral acts are always considered as independent. Repeated acts may be similar, possibly equivalent or exchangeable in terms of intentionality, but they nevertheless remain logically independent and so cannot be lumped together. Other events going on in the environment at the same time as the behavioral act of interest are independent acts not necessarily relevant to action. The aim is to gain more information concerning a single act, and not to lose the act among unrelated acts. Process analysis is person centered, not environment centered, and the unit of analysis is the behavioral act bounded solely by its own behavioral occurrence and nonoccurrence.

The differences between process and functional analysis can be illustrated by the case of a 5-year-old girl who sometimes stutters when speaking. A conventional ABC analysis looks at the situations in which the stutter occurs and notes that the girl stutters when a stranger visits, when she is anticipating something exciting, and

when people are not paying attention to her. The stutter happens at the beginning of a sentence, when her voice is loud, but not over any particular speech content, and the consequences of the stutter are that people pay more attention to her speech. One conclusion might be that the stutter is cued by an excited mood state and reinforced by attention from others, with the recommendation that a program be implemented to reinforce her differentially for not stuttering and being relaxed.

The process analysis takes one instance of stuttering and looks exactly at what the child is doing in terms of speech related actions. We look at equivalent actions where there is no stutter. We then pick out behavioral parameters, say speed of talking or rate of breathing, along which we can construct a class of speech actions. We find the values of these parameters at which the emergence of the act of stuttering is most likely when compared with other possible acts.

If a certain speech and/or breathing rate is sufficient to typify the emergence of stuttering, these can be modified by practice to give the girl greater control and flexibility of rate whenever she speaks.

The benefits of the response-centered process analysis rest principally in the ability to deal with individual actions as independent actions, and in making treatment decisions on a single instance of a behavior. In practice more empirical investigations of the action may be necessary to establish the distribution of possible responses around it, but any repetition of the act is considered as another independent behavioral act.

The approach as well as being response centered is person centered, so quite clearly the action context—the common behavioral dimension(s) along which the response class vary—are different from person to person. Also, and perhaps very differently from conventional functional analysis, situational cues, and environmental factors are not necessarily considered relevant.

One further important corollary of the response equivalence approach is that time does not necessarily become an ordering variable for behavior. The functional analysis of course must consider the temporal sequence of events to be paramount in defining process. But from the response process point of view, an act that occurs even msec prior to another act is still independent of that act. The context that defines the act is the response class of which it is a member, not the time at which it occurs.

Under certain circumstances there may be a conditional relation between independent actions. This relation is expressed additively or in combination but not as a sequential dependence. Suppose two acts I do when embarrassed are scratching my neck and coughing, and that these always occur together. Scratching and coughing are independent classes of acts. But I may combine them in this instance and talk of scratching and coughing behavior, because the response class of combined acts may give me a more precise description of what typifies this embarrassed behavior.

Similarly lighting up a cigarette and taking a puff from a cigarette are two independent classes of act. But in the context of smoking behavior one may consistently precede the other, in which case we might say not that they are correlated or associated or contingent on one another but that they are additive as part of a single continuous process.

This kind of process analysis implies a different statistical model. Because we are interested in quantifying process, we have discarded the effect model of behavior and are treating behavior as an independent variable; in so doing we have stepped outside the constraints of the usual hypothetico-deductive effect statistical models where behavior is a dependent variable.

The response process approach defines behavior by describing the context in which it occurs. The descriptive techniques of the French statistical school are therefore more appropriate, since their approach is to fit models to the data rather than data to models, and they make no hypothetical probabilistic assumptions about data.

The discussion by Rouanet, Bernard, and Lecoutye (in press) of typicality, De Finetti's (1972) concept of equivalence or exchangeability, and the correspondence analysis of Benzecri (1973) are particularly relevant to the present discussion.

However, the present chapter concentrates on presenting the logic of the process approach and statistical considerations are discussed in detail elsewhere (O'Connor, 1987).

The steps in response process analysis can be summarized as follows:

1. Detail what actions the person is actually doing during a behavior.
2. Consider a class of actions that are as behaviorally equivalent as possible to the actions in the behavior of interest.
3. Measure whatever behavioral parameters the class of actions has in common.
4. Decide on the parameters that typify the behavior of interest rather than other equivalent behaviors.

The positions of each of these equivalent behaviors can be plotted along their common behavioral dimension on either side of the value that best typifies the behavior in question. This constructs an action space around the behavior of interest. This action space also forms the basis for a likelihood distribution from which to compute the likely emergence of the behavior given the knowledge of its behavioral context so far.

CLINICAL EXAMPLE OF PROCESS ANALYSIS

As a clinical example let us take the specific case of a man suffering a panic attack in a waiting room. He is sitting in the room and a feeling of unease comes upon him. He starts to sweat, becomes fidgety, and experiences breathing difficulties. Eventually he rushes out of the door into the open air where the feeling subsides.

The specific problem behavior is the panic behavior. This is what the patient defines as the problem and as behaviorists we accept this as the problem. But in the response process version of functional analysis we look specifically at what the person is doing within the limits of this behavior.

The man is for example sitting down on a chair. This particular act of sitting down is unique in its spatio-temporal characteristics but from a motoric point of view it can be considered one of a class of sitting behaviors. The man also sits on the bus, sits watching TV, sits in the pub, sits watching football, etc. Now in not all of these sitting situations does he experience panic, so that simply knowing whether

he is sitting or not does not guarantee knowledge that he will panic or not panic. But besides sitting he is also waiting, and waiting silently. If we construct a response class of sitting waiting silently, we find that panic is likely to emerge more often during this class of response processes than when he is sitting and not waiting silently.

Having identified the general properties of the response class relevant to panic we can then look for other properties accompanying the response to enable us to typify the panic behavior with greater precision. The man may be sitting uncomfortably, he may be sitting in a particular posture, he may be looking in one direction, or thinking about a particular topic. All of these aspects form the basis of independent response classes in their own right. We may include stimulus characteristics as part of our profile, but stimuli are relevant only to the extent that they have implications for response processes, and are centered on response processes.

In Figure 4 we have constructed a response process profile of combined acts that optimally typifies the response background from which panic emerges given information so far. Here we have moved from the nominal binary classification of response classes according to presence of panic to an ordinal ranking of response profiles according to how likely panic is to emerge given a certain combination.

The process dimension goes from "sitting and waiting and not communicating with strangers" to "sitting and communicating with friends." The plot represents a two-space correspondence analysis, but we have derived it by considering a class analysis of actual actions of the person and not by imposing statistical models.

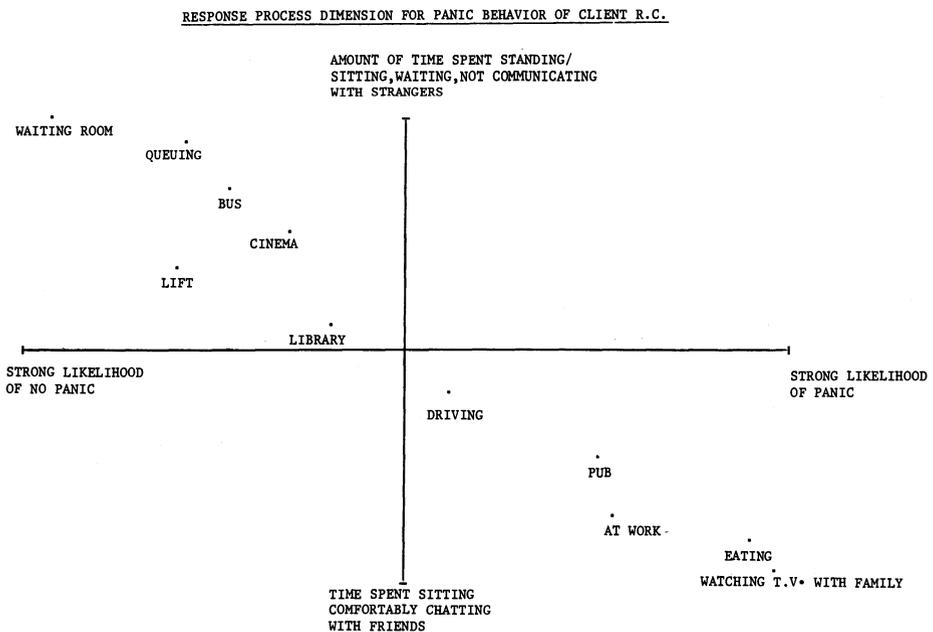


FIGURE 4. A Response Process Dimension has been constructed from the acts that a person actually does during a panic attack. The dimension represents the optimal ordering of acts according to their presence or absence during panic.

Concentrating solely on the actions that take place in a behavior, we have built up a response process definition of behavior that does not rely on commonsense assumptions and yet provides a logic for clinical-behavioral strategies in managing the problem.

Increasingly, as we saw in earlier sections, behaviorists have sought to understand the processes of response apart from sensory models but such understanding has been tied by a logic that implies a functional dependence between response and stimulus. Process analysis goes some way to relieving this theoretical burden.

CONCLUDING REMARKS

We began with the problem that in clinical practice behaviorists rely heavily on cultural assumptions in their classification of behavior. The boundaries of behavior are imputed on the basis of social constructs and ethical presumptions on conduct, not on the basis of an empirical logic that identifies behavior according to consistent response criteria.

We have discussed the insufficiencies in a classical functional definition of behavior that assumes that behavior is in some way a function of sensory processes, and suggest that behavior is better viewed as a product of response processes operating independently of sensory events. Behavior can be seen as a series of self-regulated response acts and is most likely to take the form of relational organization similar to that which controls motor action. All actions of a person at any given time are to be considered in holistic relation to a centrally planned intention. A given behavior is guided by a single future plan, not driven by a multitude of past events.

But the relational model can only offer a viable definition of behavior in response process terms if it adopts a relativist rather than a realist approach to action. An action is defined in the context of a class of response acts that are equivalent to it in terms of process rather than according to fixed sensory space-time coordinates. If behavior is not defined relative to response processes, response activity cannot be considered independent, and hence cannot reveal the structure of its own process.

A crucial and clearly controversial point in the methodology of process versus functional analysis is that the meaning of behavior is already constituted in the response processes from which it emerges. Hence the task of process analysis is simply to reveal descriptively the context of response that optimally accounts for the emergence of behavior.

Using statistical methods derived from the French school of descriptive statistics we arrive at a process dimension that lists along its axes the specific actions that optimally typify what a person actually does during various degrees of presence or absence of the problem behavior.

Experimental work could obtain more information on the specific response actions in any behavior and so add precision to the process dimension. The actions investigated would by definition be clinically relevant actions and would not depend on specified laboratory conditions for their emergence. This clearly aids generalization and congruence between clinical and experimental work.

A final point concerns ethical issues and the process approach. By not imposing sensory or cultural models on behavior this approach is able to bracket the subcultural context that may otherwise guide behavior therapy applications.

Take for example the case of a teacher who approaches a behavior therapist because she is annoyed by the chatter of her class of children as they queue up outside their room. She wants the children's behavior modified. A conventional stimulus-based functional analysis would identify the children's talking as the basis for annoyance and might embark on an operant program to modify the talking at the teacher's request (viz. Owens & Ashcroft, 1982).

The process approach considers "being annoyed by talking" as an act rather than a reaction, and seeks to construct a profile of actions of teacher and children that optimically typifies the emergence of annoyance. We may find for example that "talking loudly" as a class of response act does not typify the presence or absence of annoyance, whereas "talking loudly and being outside the room" is a response class that better typifies annoyance. But the children talking outside the room and the teacher thinking alone in the room even better typifies the behavioral processes through which annoyance emerges. In fact, the response class of the "teacher being alone in the room and thinking of a certain topic" may be sufficient process dimension alone to guarantee the emergence of annoyance by noise whatever the source. The teacher then may be phonophobic for certain emotionally selective sounds at certain times. The point is that by considering all acts involved in a behavior as equal co-respondents in defining the behavior, the process approach preempts ethically relative definitions of cause and function.

The ethical issue of who should control the behavior therapy is traditionally viewed as a worrisome but inevitable consequence of the essentially amoral character of behavioral technology (McFall, 1979). Perhaps on the contrary it is the lack of behavioral logic in defining behavior that has led to difficulty for practitioners in resisting normative definitions of behavior.

REFERENCES

- Beck, A. T. (1976). *Cognitive Therapy and Emotional Disorders*. New York: International University Press.
- Bernstein, A. S. (1979). The Orienting response as novelty and significance detector: Reply to O'Gorman. *Psychophysiology*, 16, 263-273.
- Benzecri, J. P. (1973). *L'analyse des données*. Paris: Dunod.
- Borkovec, T. D. (1979). Incubation and the relevance of functional CS exposure. *The Behavioral and Brain Sciences*, 2, 168-169.
- Coyne, J. C. (1982). A critique of cognitions as causal entities with particular reference to depression. *Cognitive Research and Therapy*, 6(1), 3-15.
- Cullen, C. (1983). Implications of functional analysis. *British Journal of Clinical Psychology*, 23, 53-55.
- Duffy, E. (1957). The psychological significance of the concept of "arousal" or "activation". *Psychological Review*, 64, 265-275.
- Erwin, E. (1978). *Behavior Therapy: Scientific, philosophical and moral foundations*. Cambridge: Cambridge University Press.
- Eyberg, S. M. & Robinson, E. A. (1985). Conduct problem behavior. Standardisation of a behavioral rating scale with adolescents. *Journal of Child Psychiatry*, 24, 51-57.
- Eysenck, H. J. (1967). *Biological basis of personality*. Springfield, IL: Charles C Thomas.
- Eysenck, H. J. (1980). A unified theory of psychotherapy, behaviour therapy and spontaneous remission. *Zeitschrift für Psychologie*, 188, 43-56.

- Eysenck, H. J. (1981). Behavioral therapy and the conditioning model of neurosis. *International Journal of Psychology*, 16, 343-370.
- Fernald, C. D., & Gettys, L. (1980). Diagnostic labels of perceptions of childrens' behavior. *Journal of Clinical Child Psychology*, 9, (3), 229-232.
- Fisher, D. L. (1982). Limited channel models of automatic detection capacity and scanning in visual search. *Psychological Review*, 89, 662-692.
- Freeman, G. L. (1948). *The Energetics of Human Behavior*. New York: Cornell University Press.
- Gale, A. (1981). EEG studies of extraversion-introversion. What's the next step. In R. Lynn (Ed.), *Dimensions of personality* (pp. 181-208). Oxford: Pergamon Press.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston, MA: Houghton Mifflin.
- Gibson, J. J. (1982). Notes on action. In E. Reed & R. Jones (Eds.), *Reasons for realism: Selected essays of J. J. Gibson*. Hillsdale, NJ: Erlbaum.
- Grant, D. A. (1964). Classical and operant conditioning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press.
- Gray, J. A. (1970). The physiological basis of introversion-extroversion. *Behavior Research and Therapy*, 8, 249-266.
- Green, D. M., & Birdsall, T. G. (1978). Detection and recognition. *Psychological Review*, 85(3), 192-206.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty. Heuristics and biases*. Cambridge: Cambridge University Press.
- James, W. (1970). *Pragmatism*. New York: Washington Square Press.
- Jones, B., & Hulme, M. R. (1976). Evidence for an outflow theory of skill. *Acta Psychologica*, 40, 49-56.
- Jung, J. (1982). *The experimenters challenge*. New York: Macmillan.
- Laming, D. (1985). Some principles of sensory analysis. *Psychological Review*, 92(4), 462-483.
- Lang, P. (1980). Emotional imagery: Conceptual structure and pattern of somatovisceral response. *Psychophysiology*, 17(2), 179-192.
- Liberman, A. H., & Mattingly, I. G. (1985). Motor theory of speech perception revised. *Cognition*, 21 (1), 1-36.
- Loveless, N. E. (1979). Event related slow potentials of the brain as expressions of orienting function. In H. D. Kimmel, E. H. van Olst, J. F. Orlebeke (Ed.), *Orienting reflex in humans*. (pp. 77-100). Hillsdale, NJ: Erlbaum.
- Markham, K. A. (1979). EEG measures of the person and person predicates. *Bulletin of the British Psychological Society*, 28, 51-54.
- Martin, I., & Levey, A. (1969). *The genesis of the classical conditioned response*. London: Pergamon Press.
- McFall, R. (1979). Behavioral training: A skill acquisition approach to clinical problems. In J. T. Spence, R. C. Carson, & J. W. Thibaut (Eds.), *Behavioral approaches to therapy*. (pp. 227-260). Morristown, NJ: General Learning Press.
- McGuigan, F. (1978). *Cognitive psychophysiology*. New Jersey: Prentice-Hall.
- Meichenbaum, D. (1978). *Cognitive behavior modification: An integrative approach*. New York: Plenum Press.
- Munsterberg, H. (1899). The physiological basis of mental life. *Science*, 9, 442-447.
- Newell, K. M. (1978). Some issues on action plans. In G. E. Stelmach (Ed.), *Information processing in motor control and learning*. New York: Academic Press.
- O'Connor, K. P. (in press). Data analysis in individual difference research. In A. Gale & M. Eysenck (Eds.), *Handbook of individual differences*. Chichester: Wiley.
- Owens, R. G., & Ashcroft, J. B. (1982). Functional analysis in applied psychology. *British Journal of Clinical Psychology*, 21, 181-189.
- Papakastopoulos, D. (1980). A no stimulus, no response event related potential of the human cortex. *EEG Clinical neurophysiology*, 48, 622-638.
- Parasuraman, R., & Abeatty, J. (1980). Brain events underlying detection and recognition of weak sensory signals. *Science*, 210, 80-83.
- Perkins, C. C., Jr. (1947). The relation of secondary reward to gradients of reinforcement. *Journal of Experimental Psychology*, 37, 377-392.
- Prokasy, W. F. (1965). Classical eyelid conditioning. Experimental operations, task demands and response shaping. In W. F. Prokasy (Ed.), *Classical conditioning*. New York: Appleton Century Crofts.
- Rouanet, H., Bernard, J. M., & Lecoutre, B. (in press). Non probabilistic statistical inference: A set theoretical approach. *American Statistician*.

- Sabat, S. R. (1979). Selective attention, channel capacity and the averaged evoked response in human subjects. *Neuropsychologia*, *17*, 103–106.
- Savage, L. J. (1972). *The foundation of statistics*. New York: Dover.
- Schachter, S. (1971). *Emotion, obesity and crime*. New York: Academic Press.
- Siddle, D. (1985). *Orienting and habituation. Perspectives in human research*. Chichester: Wiley.
- Sperry, R. W. (1969). A modified concept of consciousness. *Psychological Review*, *76*, 532–536.
- Strelau, J. (1983). *Temperament personality activity*. London: Academic Press.
- Treisman, A. M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97–136.
- Turvey, M. T. (1977). Contrasting orientations to a theory of visual information processing. *Psychological Review*, *84*, 67–88.
- Turvey, M. T., Shaw, R. E., & Mace, W. (1978). Issues in the theory of action: Degrees of freedom, coordinate structures and coalitions. In J. Requin (Ed.), *Attention and performance, VII*. (pp. 557–595). Hillsdale, NJ: Erlbaum.
- Watson, J. B., & Raynor, R. (1920). Conditioned emotional reactions. *Journal of Experimental psychology*, *3*, 1–14.
- Weimer, W. B. (1977). Motor theories of mind. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing: Toward an ecological psychology*. (pp. 270–311). Hillsdale, NJ: Erlbaum.

PART V

**BIOLOGICAL BASES OF
PERSONALITY AND BEHAVIOR**

CHAPTER 18

The Role of Heredity, Environment, and “Preparedness” in the Genesis of Neurosis

H. J. Eysenck

As noted in a previous chapter, Watson’s theory of neurosis is untenable in its original form, and inevitably the 65 years that have elapsed since its publication have unearthed a great deal of evidence to suggest ways in which the theory could be improved, and in part altered. Previous chapters have dealt with changes in the interpretation of conditioning, incorporating cognitive mechanisms, the development of the notion of “incubation of anxiety,” and the shift from Pavlovian A to Pavlovian B conditioning; and the general shift away from S-R to S-S interpretations of conditioning phenomena. In this chapter we will deal with a rather different set of assumptions made by Watson, but clearly erroneous, and shown to be untenable by recent work. These studies deal with the alleged equipotentiality of stimuli, that is, the notion that from the point of view of conditioning all stimuli, however artificial, may be considered equally likely to produce conditioned responses when paired with the UCS; and the more general question of the preparedness of certain stimuli to become associated with UCSs. These problems are considered within the general framework of the relative importance of genetic and environmental factors in predisposing a person to develop neurotic disorders.

Watson, like most behaviorists, stressed environmental causes of behavior, and played down as much as possible the importance of genetic causes. He was not entirely consistent in this, however; Watson and Rayner (1920) speculated that the long continuance of little Albert’s conditioned fear reactions to rats might have been due in part to some kind of “constitutional inferiority,” and might not have been observed in other children presumably not suffering from such an inferiority. The term itself is not defined, but clearly Watson did consider the possibility that individual differences played a large part in the development of conditioned fear reactions,

leading to neurotic disorders and phobias, and that these might be of genetic origin. However, his behaviorist successors have certainly paid little attention to this possibility, and to this day adopt an almost 100% environmentalism.

Psychiatry, largely under Freudian influence, showed an equally clear disregard of genetic factors (not actually shared by Freud). Here, for instance, is a quotation from a widely used textbook by Redlich and Freedman (1966) entitled *The Theory and Practice of Psychiatry*. They only make one comment on the importance of genetic factors in mental disorders: "The importance of inherited characteristics in neuroses and sociopathies is no longer asserted except by Hans J. Eysenck and D. B. Prell" (p. 176). The reference is to a paper by Eysenck and Prell (1951), who provided evidence of a high degree of heritability in neuroticism, using pairs of MZ and DZ twins; a little later the demonstration was extended to the heritability of extraversion-introversion by Eysenck (1956).

The statement itself was of course completely untrue; Eysenck (1967) has quoted ample evidence even from these early days for genetic effects in the causation of neurotic disorders, and many other summaries can be quoted in support (e.g., Miner, 1973; Roubertoux & Carlier, 1972, 1973; Slater & Shields, 1969). The quotation merely illustrates the deliberate refusal of many leading psychiatrists and psychologists to look at the evidence, and to acknowledge the importance of genetic factors.

Any discussion of heritability in this context is handicapped, not so much by the fact that psychologists are normally not exposed to any systematic teaching of modern genetic theory and practice, but rather because they share a number of erroneous assumptions that make it difficult to discuss such concepts as heritability, interaction, and environmental variance. The first misconception is that modern behavioral genetics is only concerned with genetic causes, and hence almost by definition biased in favor of finding such causes. This is quite incorrect. Modern genetic theory, as Fulker (1981) makes clear, is concerned with what he calls "the genetic and environmental architecture" of the causal factors underlying the phenotypic observations that constitute the raw material of psychological science. Being concerned with the breakdown of the phenotypic variance, genetic theory cannot arbitrarily restrict itself to the genetic portion of the variance; it must inevitably consider this as a portion of the total phenotypic variance, and hence assess the contribution of environmental variance also.

The next point to be noted is that genetic and environmental variance are nowadays broken down into several distinct components, so that the ascertainment of heritability is only one, and probably not the most important, task of modern behavioral genetics. Total genetic variance (V_G) is made up of additive genetic variance, that is, the simple additive action of separate genes making for high or low intelligence, neuroticism, extraversion, or whatever; this is denoted as V_A . Next we have nonadditive genetic variance due to dominance at the same gene loci (V_D), and nonadditive genetic variance due to interaction between different gene loci, called *epistasis* (V_{EP}). Finally, there is genetic variance due to *assortative mating*, that is, the increment in total variance attributable to degree of genetic resemblance between mates in the characteristic in question (V_{AM}).

As regards the environmental variance, it is useful to decompose it into environmental variance between families (V_{EB}), and environmental variance within families (V_{EW}). The former refers to systematic environmental differences between families

that make for differences between offspring on the trait in question, but do not make for differences among offspring reared together in the same family. Opposed to this there are differential and environmental influences within families that make for differences among offspring reared together in the same family.

We can now define heritability, which is given by the formula:

$$h^2 = \frac{V_G}{V_P}$$

where V_P is the phenotypic or total variance of the trait or behavior in question. The phenotypic variance is made up as follows:

$$V_P = V_G + V_E + V_{GE} + \text{Cov}GE + V_e$$

where V_G and V_E refer to the genetic variance and the (additive) environmental variance that is independent of the genotype, respectively. V_{GE} refers to variance due to interaction, that is, nonadditive effects of genotypes and environments, and $\text{Cov}GE$ refers to the covariance of genotypes and environments, whereas V_e refers to the error variance due to unreliability of measurement.

It is important to distinguish the differences between the two interaction terms. V_{GE} means that different genotypes may respond differently to the same environmental effect. Thus if coaching on an IQ test, say, raises the IQ of every genotype subjected to it by 10 points, the environmental effect is said to be *additive*, and the variance contributed by such an environmental effect is included in V_E . If, on the other hand, administration of a drug like glutamic acid causes genotypes with lower IQ to gain more IQ points than genotypes with average IQ, and lead to no gain at all among those with superior IQs, then the environmental change *interacts* with genotypes to produce different phenotypic effects in different subgroups. This source of variance is called V_{GE} . The covariance between genotypes and environments, $\text{Cov}GE$, arises when genotypic and environmental effects are *correlated* in the population. Thus if children with genotypes for high intelligence are also reared in homes with superior environmental advantages for intellectual development, such covariance arises. Some part of it is of course itself the product of the genotype, as when an intellectually gifted child spontaneously spends much time in reading or other intellectual activities.

Heritability can, in fact, be defined along different lines. *Narrow* heritability is the proportion additive genetic variance is of total phenotypic variance (V_A/VP); broad heritability has already been defined as V_G/VP . In the usual formulae, the error variance (V_e) is included with the environmental variance, and hence the estimate of genetic variance is too low, and a suitable correction should be made. This underestimation can be quite serious, and in what follows attention will be drawn to this point.

Modern methods of analysis, using data from identical twins brought up in separation, comparisons between MZ and DZ twins, studies of adopted children, familial intercorrelations, genetic regression to the mean, inbreeding and heterosis effects, and many other methods enable us to give estimates of the different portions of these formulae (Fulker & Simmel, 1983; Mather & Jinks, 1971). It is also possible to assess the power of these methods (Martin, Eaves, Kearsley, & Davies, 1978), and

to estimate the numbers of twin pairs, say, required to give a particular set of fiducial limits for one's estimates. This is obviously not the place to go into these technicalities, and the reader must be referred to the sources cited.

Another frequent error in this field is to regard estimated heritabilities as applying to individuals. As will be clear from the fact that we are using analysis of variance, heritabilities are population estimates; in other words, they apply to groups of people, say British people living in Great Britain and born between, say, 1930 and 1960. The error of arguing as if heritability estimates pertained to individuals is clearly brought out by an argument originally brought forward by Donald Hebb, who suggested that trying to estimate the relative importance of genetic and environmental factors was as silly as to try and say whether length or width of a field was more important in defining its area. The single field of course has no variance, and consequently the comparison does not apply; if we asked whether among a hundred fields length or width was more important, the matter could be easily subjected to a statistical test.

The fact that we are dealing with population estimates also serves to clarify another frequent error that is made by writers in this field. Heritabilities are not given once and for all, but apply to a given population at a given time. Subdividing samples of twins in Norway into age groups widely separated from each other, and looking at heritability of scholastic achievement, Health *et al.* (1985) found that as expected heritabilities were highest for the youngest age group, lowest for the oldest, and intermediate for the middle-age group. The obvious explanation is that, increasingly greater equality of education in recent years reduces the environmental component over time.

This example also illustrates another common error, namely that any trait or characteristic that is at least in part inherited is thereby fixed for all eternity. This is clearly untrue; changing environments will change heritabilities. It is easy to imagine that genetic causes exert a completely deterministic effect on individual behavior, but this clearly is not so. To understand that genetic and environmental factors always work in interaction, in very complex ways, is the beginning of wisdom in approaching the whole topic of behavioral genetics.

A slightly fictitious example may make this point clear. At the moment, in conditions of adequate nourishment, the size, the shape, and consistency of the female bosom is determined very largely by genetic factors, and exercise, massage, etc., have little control over it. However, recent advances in hormonal treatment, plastic surgery, and silicone injections have altered the situation to such an extent that it is quite conceivable that in 50 years time, in California, genetic factors will play very little role in determining the size, shape, and consistency of the female bosom. In a similar way, it may be suggested, the introduction of behavior therapy may have altered the strong genetic determination of neurotic disorders.

Turning now to a substantive account of work on the genetics of personality, we may note that the origin of the belief that genetic factors played little part in personality can be found in the work of Newman, Freeman, and Holzinger (1937), who published a study on twins that has often been cited, but that is subject to many serious criticisms (Eysenck, 1967). The personality tests used were inappropriate to the age group tested, they were unreliable and probably invalid, and the conclusions

drawn by the authors from the more valid and reliable tests, such as a neuroticism questionnaire, were counter to their own findings. On the Woodworth-Mathews Inventory, for instance, MZ twins obtained an intraclass correlation of .56, whereas DZ twins obtained one of 0.37; for MZ twins brought up in separation the intraclass correlation was .58. This suggests that heritability is between .38 and .58, surely not all that low, particularly when we correct the observed figures for attenuation, which would bring them into the .5 to .7 region. Other studies surveyed by Eysenck (1967) also indicate quite clearly the importance of genetic factors as far as neuroticism is concerned, as does the work of Eysenck and Prell (1951).

All these studies suffer from two defects. In the first place, the number of twins studied was too small to give any information other than that genetic factors are probably involved, but not permitting any very clear quantitative estimation, or any study of the different types of genetic and environmental factors entering into the equation. Secondly, methods of analysis had not yet developed beyond the very elementary stage, and it is only recently that methods have become available to do a proper genetic analysis of the large-scale twin data now available.

At the point of transition stands the important work of Shields (1962), in which he used an early version of the Eysenck Personality Inventory. He studied 44 separated MZ, 44 nonseparated MZ, and 32 pairs of DZ twins of which 11 had been brought up apart. For neuroticism, separated twins showed an intraclass correlation of .53; the nonseparated ones of .38, and the DZ twins of .11. These results are in good agreement with those originally reported by Newman, Freeman, and Holzinger, leading to similar estimates of heritability. The work was so carefully done, and the results so extensively reported, that it is still possible to reanalyse the data by more modern methods.

The modern period of investigation may be said to have begun with a series of papers by Eaves and Eysenck (1975, 1976a,b, 1977; in press). (Eaves and Young, 1981, may be consulted for a detailed account of the development of this work.) For the first time we have here large enough samples of MZ and DZ twins, both same-sex and differently sexed, to make proper calculations possible, and at the same time we have the first application of the new model-fitting methods of the Birmingham school. Model fitting proceeds by taking a very simple model, assuming, say, that all the variance is due to between-family environmental variance; when this model fails the chi-square test, we may add other factors, such as within-family environmental variance, or additive genetic variance, etc. In this model-fitting process, we can see which are the variables that improve the model significantly, and which must therefore be retained, and which are the variables that do not improve the model significantly, and must accordingly be rejected.

In our own work, we have found, as also reported in subsequent work by others (e.g., Floderus-Myrhed, Pederson, & Rasmuson, 1980; Jardine, Martin, & Henderson, 1984; Kendler, Heath, Martin, & Eaves, 1986; Loehlin & Nichols, 1976; Martin & Jardine, 1986), that there are certain regularities that appear again and again. In the first place, neuroticism (and the other major dimensions of personality, such as extraversion-introversion and psychoticism) has heritabilities around 50%, which rise to above 60% when corrected for attenuation. In the second place, there is no evidence for any important contribution by between-family environmental variance

for any of the personality variables, an important point because traditional theories in personality research refer almost exclusively to variables of precisely this type. In the third place, environmental variance is contributed almost exclusively by within-family factors. In the fourth place, errors of measurement play an important part, and because these are in the usual formulae confounded with within-family environmental variance, it is important to correct the obtained heritability coefficients for this factor. In the fifth place, there is evidence in the larger studies at least, for dominance effects in extraversion, but not in neuroticism or psychoticism. For neuroticism in particular, the genetic contribution seems to be almost entirely additive in nature. In the sixth place, there are in the larger studies important age and sex factors. In fitting a model, we start out with the assumption that the model will fit equally the older and the younger, and males and females. This assumption can be checked, and if found incorrect leads to important findings concerning age and sex variables. These we will discuss presently.

These are the major findings emerging from this very large body of work, and it is reassuring to be able to note that studies carried out by different investigators, in different countries (Great Britain, United States, Australia, Scandinavia) using different methods of analysis, different samples, and different questionnaires, give results that by and large are in astonishingly good agreement. If replication is the life blood of science, then clearly behavioral genetics, as far as personality is concerned, is in good shape.

The first set of studies to be discussed now is the one by Eaves and Eysenck (1976a,b, in press). Six basic models were fitted to the data for male and female MZ twins, male and female DZ twins, and a male-female group of DZ twins. The models were E_W only; E_W and E_B only; V_A and E_W only (the genetic model); E_W , E_B and V_A ; E_W , V_A , and V_D (dominance model); and V_A , $V_{A'}$, E_W (the competition-cooperation model) in which $V_{A'}$ represents the genotype covariance generated if the same genes have a direct effect on the trait of one twin and an indirect effect on the environment of a co-twin.

The E_W model and the E_W and E_B (the environmental model) are soundly rejected for both extraversion and neuroticism. The third model, which assumes only additive genetic effects and within-family environmental causes, gives a very good fit for both. Little improvement follows the addition of the family environment (Model 4), dominance (Model 5), or competition (Model 6). Thus, it is fairly certain that the family environment does not contribute significantly to variation in extraversion or neuroticism. The possibility exists that some dominance effects may have been missed in this study because the sample size is not large enough (some 500 pairs of twins) to detect this unambiguously. However, on the whole the assessment of the London data for extraversion and neuroticism lead to the proposition that additive gene action, within-family environmental effects, and little or no effect of the family environment provide an adequate description of the data.

Similar results to those of London study were reached by Eaves and Young (1981) in a reanalysis of the extensive data from the National Merit Twin study (Loehlin & Nichols, 1976). As in the London study, raw scores were transformed because of skew. Unfortunately, in this study there were no data on unlike-sex twin pairs. This meant that, although sexes can be compared for the scale on which genes

and environmental effects are expressed, it is impossible to judge whether such differences are due to effects of the same genes in both sexes or not. In general, studies that retain unlike-sex twin pairs are to be preferred when there is little knowledge about the significance of sex differences in the determination of individual differences. Because the sample was of much more uniform age than the London sample, no age correction was undertaken.

The results of the model fitting are strikingly similar for extraversion and neuroticism, and confirm the findings for the British sample. Both extraversion and neuroticism are consistent over sexes with respect to the causes of variation, no matter which model is assumed at the outset. It is also clear that the fit of the simple genotype-environment model is far superior to that which assumes no genetic variation but allows for family environment.

Wherever we examine the results in the tables we shall find support for our conviction that the effects on the family environment are small by comparison with the effects of genotype or within-family environment. . . . it would thus appear that two large bodies of personality data, one British and one from the United States, agree broadly in showing that while effects of additive gene action can be demonstrated with reasonable reliability there is little support for the view that social factors are affecting personality in so far as these depend either on the phenotypes of the individuals themselves or their parents. (p. 148)

Sample sizes in these two studies were between five and eight hundred pairs of twins; a third and still larger body of data relating to personality has been published by Floderus-Myrhed, Pederson, and Rasmuson (1980). The authors obtained an unselected sample of 12,898 twin pairs of like sex. Because of the enormous size of the sample they were able to divide their data into three age cohorts, and report the mean squares from the analysis of variance separately by sex. Both extraversion and neuroticism were assessed by 12-item untransformed scales taken from the Eysenck Personality Inventory. Model-fitting techniques were applied by Eaves and Young (1981) to the authors' own data summaries, and the results, although in essence bearing out the conclusions from the London and the United States studies, add important new insights because of the reduced size of the fiduciary limits. Again we find that environmental models are completely inadequate, and that the genetic model of additive genetic variance plus within-family environmental variance gives the most adequate account. However, these more extensive data question the validity of the additive model for extraversion, suggesting the incorporation of dominance in the model, and raise the possibility of age and sex-dependence for gene action. The most marked trend as far as age is concerned is toward environmental variance with decreasing age. It is also clear that the contribution of genetic differences is greater in males than in females. The dominance components for extraversion shows an increase with age in females and a decrease in males.

In view of these several interactions it is necessary to quote separate heritability estimates for each group. These are given in Table 1. It will be seen that the estimated contribution of genetic factors shows a similar trend for both traits. Such trends, of course, can arise from changes in the genetic variance, the environmental variance, or both, but obviously heritability is stronger for females than for males, and stronger in the younger than the older group.

In addition to the greater influence of genetic factors on females, we should note that in this analysis there is a significant improvement when the parameters are

TABLE 1. Heritability Estimates for Swedish Data

	Cohort	Neuroticism	Extraversion
Male	1926–1935	0.38	0.45
Male	1936–1945	0.39	0.38
Male	1946–1958	0.53	0.51
Female	1926–1935	0.40	0.50
Female	1936–1945	0.51	0.49
Female	1946–1958	0.61	0.57

*From Eaves and Young (1981).

allowed to vary independently over sexes, implying that the magnitudes of the components are different between males and females. The model fitting requires a similar conclusion with respect to the effects of age. Thus we must conclude that both age and sex, when considered separately, affect the expression of genetic and environmental differences in extraversion and neuroticism. It is only when separate estimates are obtained for each sex and age combination that the model really fits this large section of data. Thus the Swedish data reveal an interaction between age and sex and the expression of genetic and environmental influences for the personality measurements, and the data require six separate sets of parameter estimates in order to obtain a model that accounts adequately for the entire data set. The introduction of dominance into the model does nothing for neuroticism, but its addition to the extraversion model leads to a highly significant improvement in fit. Thus the Swedish data give us important additional information, as well as confirming the importance of genetic and within-family environmental variance.

A fourth large-scale study has been reported by Martin and Jardine (1986) and Jardine, Martin, and Henderson (1984). They report results on 3,810 pairs of twins of the Australian Twin Registry. These twins were administered the Eysenck Personality Questionnaire (Eysenck & Eysenck, 1975), and the anxiety and depression scale of the Delusions-Symptoms-States Inventory of Bedford, Foulds, and Sheffield (1976). This study is of particular interest because it looks at the genetic and environmental causes of anxiety and depression, as well as neuroticism, testing in this way Eysenck's theory that these are two of the component features of neuroticism, which however also contains other components, such as guilt, worry, etc. Figure 1 shows roughly the postulated relationships.

The raw data were scaled, as is usual in genetic analyses, so that genetic and environmental effects are additive. The results of the analyses are very similar in principle to those already discussed, with the only model to emerge with any credit being the genetic model, that is, $V_A + E_M$. For extraversion, we obtained significant contribution by V_D , that is, the dominance component already revealed in the Swedish data is again brought out quite clearly.

Table 2 shows the sources of variance, including V_e , bringing out for the first time the importance of error variance. This reveals quite clearly that the genetic contribution listed under V_A is very much underestimated when errors are left out of account; it will be seen that when we look at the true variance of the scales used,

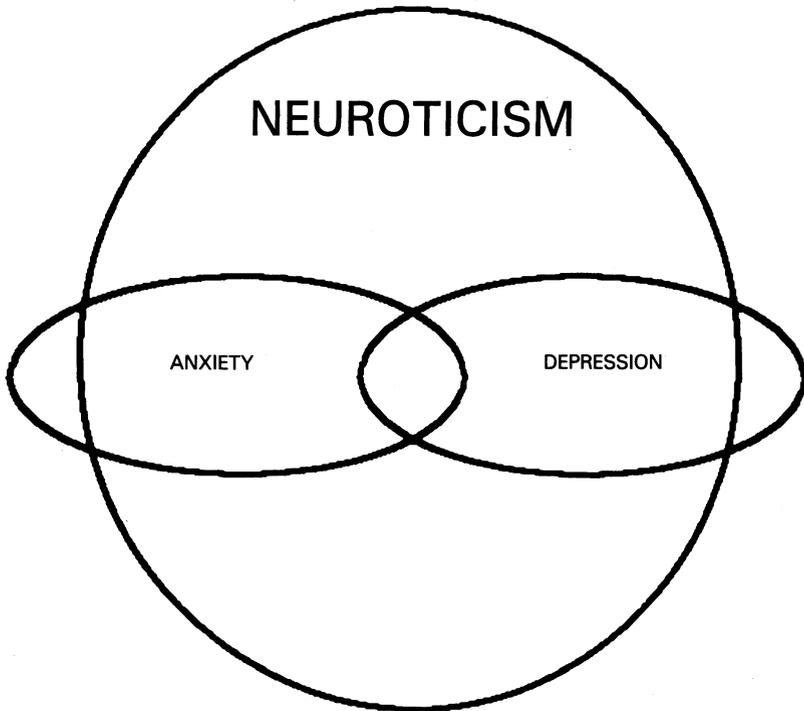


FIGURE 1. Postulated relationship between neuroticism, anxiety, and depression.

TABLE 2. Sources of Variance for Anxiety, Depression, and Neuroticism

	E_w		V_A		V_e	
	Male	Female	Male	Female	Male	Female
Anxiety	22	23	33	39	45	38
Depression	38	30	33	37	29	33
Neuroticism	37	36	45	51	18	13

Note: E_w (within family environmental variance); V_A (additive genetic variance); V_e (error variance). (Martin & Jardine, 1986)

only V_A is well above the 50% level. It would also seem that the genetic contribution is greater for females than for males, although the differences are not very large.

We now turn to the major part of the analysis, namely the causes of covariation between anxiety, depression, and neuroticism. The results of the analysis suggest that genetic variation in the symptoms of anxiety and depression is largely dependent on the effects of the same genes that determine variation in the trait of neuroticism. However, there is still substantial specific genetic covariance for neuroticism, suggesting that this may be manifested by additional components to anxiety and depression, as shown in Figure 1. There is also some evidence to indicate that there may be systematic environmental experiences influencing the trait of neuroticism that do not influence the symptoms measured, and are additional to them.

It is possible to calculate the genetic and environmental correlations of the three variables in question. In both sexes, genetic correlations are much higher (around 0.8) than the corresponding environmental correlations (around 0.4), and are similar for the three variables. Although a distinction has been made between personality traits and states (Foulds, 1965, 1974), for the neurotic symptoms measured in this study there is good evidence for a common genetic and within-family environmental basis.

It should be noted that there are also substantial genetic effects on neuroticism (16% of the total in females, 12% in males) that are independent of the two symptoms of anxiety and depression. Although specific genetic variance is a small proportion of the total for depression (6% in females, 9% in males) it is possible that this fraction estimates the contribution made in this sample by the major gene polymorphisms that are alleged to predispose to major depression (Comings, 1979; Weitkamp, Stanger, Persad, Flood, & Guttormsen, 1981). On the other hand, the genetic factor variance (19% in females, 23% in males) may be regarded as a fraction contributing to neurotic or minor depression.

Martin and Jardine (1986) summarize their results as follows:

The data suggested that population variance in these measures is due only to additive genetic effects and the influence of environmental factors which are unique to the individual. Both symptoms appear to be influenced largely by the same genes in both sexes, but have greater effect in females than males. Environmental variance for depression is also greater in females, a result found previously by Eaves and Young (1981). We found no evidence for the importance of environmental influences shared by members of the same family, effects, such as social class and parental treatment. Workers who postulated that early environmental experiences are a major influence on anxiety and depression in adulthood (Parker, 1979, 1981a, 1981b) must recognize that such experiences are not necessarily shared by co-twins; experience from parents is more likely to be a function of the child's genotype than of the family environment. (p. 41)

Martin and Jardine make the important point that

the detection of considerable genetical non-additivity for extraversion contrasts well with the lack of evidence for dominance variance affecting neuroticism, and reinforces the view that these two traits are not only statistically independent but also quite independent in fundamental biological aspects. (p. 42)

As they point out, this finding speaks strongly against Gray's (1970) argument that a 45% rotation of Eysenck's extraversion and neuroticism dimensions is justified on several biological grounds. The genetical analysis ascribes quite different origins to the genetic variation for E and N , and rotation would obscure this distinction.

Even more specific than these data has been the analysis published by Kendler, Heath, Martin, and Eaves (1986), which looks separately at each of the seven anxiety and seven depression items in the two scales. By and large these were also found to fit the model, and demonstrated therefore that for the majority of the symptoms, the $V_A + E_W$ model with sex-dependent thresholds provided a good fit. Details about the few departures from this generalization will be found in the article mentioned.

A book by Eaves and Eysenck (in press) on the genetics of personality also makes a detailed examination of the genetical analysis of individual items on the *EPQ* neuroticism and extraversion scales, but interesting though the results are, they

are not strictly relevant to our topic. Quite different is the position of Torgersen's (1979) analysis of the nature and origin of common phobic fears. He analyzed the genetic determination of five factors isolated from an analysis of a 38-item questionnaire of phobic fears, the factors relating to separation fears (Factor 1), animal fears (Factor 2), mutilation fears (Factor 3), social fears (Factor 4), and nature fears (Factor 5). Using a relatively small sample consisting of 99 same-sex pairs of twins, Torgersen found heritabilities of .23, .47, .48, .50, and .53 for the five phobic factors, but this is not the most important result of his study. Of particular interest is the question of whether all the genetic variance of these phobic fears is common, or whether specific phobic fears are genetically determined to any marked degree. In order to look at this problem, Torgersen analyzed his sample into those pairs who had the same most pronounced fears, and those who had different most pronounced fears. For MZ and DZ twins respectively, the figures were 49% and 30% for the same most pronounced fears, indicating a significant tendency for specific phobic fears to be inherited in this sample.

An interesting finding in the Torgersen paper is related to the question of whether the greater childhood similarities in MZ paternal treatment, as compared with DZ twin pairs, is responsible for greater MZ intraclass correlation. When the MZ group was separated into two groups, one with more similar and one with more dissimilar childhood environment, there were no intrapair concordance differences with respect to the strength of phobic fears between the two groups. An analysis of DZ twins pairs give the same results.

In other words, the extent of similarity in environment and childhood is unrelated to similarity in strength and content of phobic fears in both MZ and DZ twin groups and consequently cannot explain the results of this study. (p. 349)

These results are in good agreement with the reports of Loehlin and Nichols (1976), who also found that similarity of treatment played no part in greater similarity of twins for intelligence or personality. Criticisms of the twin method based on this factor of similarity of treatment may therefore be dismissed as outdated.

Other recent studies extend the investigation of genetic factors to personality traits more loosely related to neuroticism, such as empathy, aggressiveness, and assertiveness (Rushton, Fulker, Neale, Nias, & Eysenck, 1986); and impulsivity and sensation-seeking behavior (Eysenck, 1983), whereas others (e.g., Katz & McGuffin, 1986) have looked at evidence from familial investigations, for example, examining the relationship between personality factors, such as neuroticism and depression, in subjects who may have a familial vulnerability to depression (i.e., first-degree relatives of depressed patients). But although these are of interest, they are perhaps marginal to the concern of this chapter.

We may conclude from this brief survey of a vast amount of data and highly complex analyses that Watson was wise in his suspicion that little Albert (and others who retained for long periods of time conditioned fear responses in spite of opportunities for extinction) may have suffered from an "inferior constitution," which may be translated into "being endowed with a high degree of neuroticism and a high degree of introversion," the two variables predisposing a person to the development of strong emotional reactions, and enabling him to form conditioned responses quickly

and strongly (Eysenck, 1981). These findings thus integrate well with the theory outlined at the beginning of this book.

Most of the work cited so far has dealt with the genetics of neuroticism and it is logically possible that neurosis might not share the genetic determination of neuroticism. The threshold model (diathesis-stress model) of neurosis, originally introduced by Slater (1943); Slater and Slater, (1944), and in a series of studies by Eysenck, summarized in two books (Eysenck, 1947, 1952) suggests that neuroticism is the predisposing factor, neurosis developing as a consequence of stress produced by the environment. If this model were correct, then we would expect neurotic breakdown also to be affected by heredity, and reviews by Schepank (1973) and Carey and Gottesman (1981) leave little doubt that indeed various types of neuroses are strongly influenced by genetic factors. If the evidence seems less conclusive and impressive, this is because of the fact that fewer studies are available, and that the data themselves, depending as they do on psychiatric diagnoses, are notoriously unreliable. Such unreliability, by increasing the V_e factor in our formula, would inevitably increase the apparent contribution of environmental factors, and decrease that of genetic factors. Correction of the reported figures for this unreliability would almost double the true genetic contribution.

These findings also suggest that the equipotentiality of conditioned stimuli, assumed by Watson and others to be a feature of the conditioning paradigm, must in fact be rejected. If certain stimuli have a genetic basis that provokes fear reactions and anxieties, then it would be natural for stimuli associated with the fear-producing stimuli to be more readily conditioned than others that have no connection with fear-producing stimuli. The notion of belongingness, first introduced by Thorndike (1931), and recently made the subject of determined experimental investigations (e.g., Hamm & Vaitl, 1985), has led to Seligman's very influential concept of preparedness (1971).

This concept of preparedness was introduced in part because the notion of equipotentiality failed to account for major facts that had been discovered about phobias. Phobias appeared to be acquired quickly, whereas laboratory fears typically require a series of trials before they are established. Again, whereas phobias are often extremely persistent and need special curative measures to be extinguished, laboratory-conditioned fears weaken quickly when the CS is presented without reinforcement from the UCS. Another difference is that phobias are irrational in their refractoriness to verbal persuasions regarding the lack of real danger involved, whereas conditioned fear responses in humans are strongly influenced by reassuring verbal instructions. Finally, phobias are typically related to a nonarbitrary set of situations that differ markedly from those one would expect from learning theory. Phobias are seldom related to events giving rise to painful experiences in modern life, such as electrical equipment or dental treatment, but are much more frequently related to relatively rare and often harmless events and organisms. Thus, fear of snakes is about twice as prevalent as fear of dental treatment in a normal population, although presumably many more persons have had actual painful experiences in the latter context (Agras, Sylvester, & Olivean, 1969).

According to Seligman (1971), what is distinctive about the various categories of events associated with phobias is that they reflect potential dangers to the survival of prehistoric man and his mammalian ancestors, and indeed it is clear that objects

and circumstances like snakes, heights, and large open or small enclosed spaces may have involved a considerable degree of danger. Numerous factorial analyses of circumstances, events and organisms giving rise to phobic fears (Arrindell, 1980; Arrindell & van der Ende, 1986; Arrindell & Zwaan, 1985; Arrindell, Emmelkamp, & van der Ende, 1984; Dixon, de Monchaux, & Sandler, 1957; Doctor, 1982; Hafner & Ross, 1984; Hallam & Hafner, 1978; Hersen, 1973; Tasto, 1977; Wade, 1978; and Wolpe & Lange, 1964) testify to the fact that indeed the sources of phobic fear constitute a very restricted sample of such situations and organisms, and by and large support Seligman's notion that phobias are cases of biologically prepared learning. In other words, he is suggesting that humans have a biologically derived readiness to associate easily fears with typical phobic situations. It is for this reason that phobias are assumed to be rapidly acquired, resistant to extinction, insensitive to cognitive factors, and selective to stimulus situations.

There is certainly much evidence that discrimination, generalization, and categorization occur at a very sophisticated level even in such "dumb" animals as pigeons. Thus pigeons and other animals can categorize photographs or drawings as complex as those encountered in ordinary human experience (Herrnstein, 1985). Apparently animals have "natural categories" to use in this respect, and these "natural categories" are no doubt, as are phobic fears, the outcome of millions of years of natural selection. The literature is now large (Cabe, 1976; Cerella, 1979; Herrnstein, 1979, 1980, 1984; Herrnstein & Loveland, 1964; Herrnstein, Loveland, & Cable, 1976; Quine, 1969), and leaves little doubt on the point. Perception, generalization, categorization, and conditioning all proceed on a basis that is not the *tabula rasa* of Locke, but rather an inherited basis that helps to bring order into a confused environment and helps the individual organism to survive.

The work on this specific area of preparedness of phobic fears has been extended very much by the Uppsala School, and an outstanding summary of the work of this school has been given by Öhman, Dimberg, and Öst (1985). These studies compare the effect of potentially phobic CSs with nonphobic ("neutral") stimuli, using shock as the UCS and recording galvanic skin responses (GSRs) as CRs. Slides of snakes and spiders are typical of the sort of stimuli that posed threats to our ancestors, and are used as potentially phobic CSs, whereas pictures of houses, faces, circles, triangles, flowers, and mushrooms have been used as neutral stimuli. The theory tested is of course that fear responses should be more easily conditioned and more resistant to extinction when phobic rather than neutral CSs are presented.

There is good evidence that potentially phobic stimuli are reliably more resistant to extinction than responses to neutral stimuli (Hugdahl, Fredrikson, & Öhman, 1977; Öhman, Eriksson, & Olöfsson, 1975; Öhman, Eriksson, & Löfberg, 1975; Öhman, Fredrikson, Hugdahl, & Rimmö, 1976; but see McNally & Foa, 1986), but superior acquisition to the phobic stimuli has proved elusive, except in a study by Hugdahl, Fredrikson, and Öhman (1977). More recent work by Hugdahl and Kärker (1981), Hugdahl and Öhman (1980), and the review in Öhman (1979) leave little doubt on the subject.

Failures to replicate in other laboratories have been reported by Cook (1981) and Hodes (1981); however, Cook (1983) has convincingly shown that these failures were caused by the choice of aversive noise as the UCS, instead of the more usual

electric shock. He argued that from an evolutionary point of view, a tactile component might be essential in the UCS because when predators produce pain they are more likely to do this through insults to the skin. Thus the notion of preparedness might have to be extended to the UCS as well as to the CS.

The Uppsala work was extended to the acquisition of fears through symbolically transferred information (Rachman, 1977). Hugdahl and Öhman (1977) developed an experimental analogue for that phenomenon by threatening their subjects with an electric shock after either one of two animals stimuli (snakes and spiders), or two neutral stimuli (flowers and mushrooms). They found significantly better differential responding between the two former stimuli than for the latter, an effect also observed by Hugdahl (1978), who also found that instructions were as efficient as electric shock UCSs inducing skin conductance (SCR) responding.

Data involving heart rate conditioning have been less decisive than those from SCR conditioning, but on the whole the results do not contradict the hypothesis (Öhman, Dimberg, & Öst, 1985). There is, however, some interesting indication in these data that responses acquired through instructions differ from those acquired through direct conditioning.

The Uppsala group also tested another prediction from the preparedness theory, namely that responses to potentially phobic stimuli should show evidence of autonomic rather than voluntary control, implying that once acquired these responses should not be affected by conscious intentions as manipulated through verbal instructions to the subjects. Results have been favorable in the Uppsala laboratories (see Öhman & Hugdahl, 1979, for a review), but others (e.g., Cook, 1983; Dawson & Schell, 1985), have been unable to replicate the results. This is an important area of research, and more definitive answers to the questions should soon be forthcoming.

Delprato (1980) pointed out that the work of Rachman and Seligman (1976) and De Silva, Rachman, and Seligman (1977) threw some doubts on the preparedness hypothesis. In the former studies patients with "unprepared" phobias did not, as expected in terms of the theory, prove very responsive to treatment. In the second study, a retrospective analysis of a large number of phobic and obsessional cases treated at the Maudsley Hospital over a 5-year period produced as the most important finding the fact that preparedness was unrelated to the ease of acquisition, and to therapeutic outcome. This is suggested to pose serious problems for the preparedness concept.

One's evaluation of these findings depends a good deal on a number of questions that arise. Öst and Hugdahl (1981), in a study of the causes of phobic responses, reported that just under 50% of their sample attributed the onset of their animal fears to a traumatic conditioning episode, whereas slightly less than a third had vicarious conditioning experiences, and about one sixth reported instructions and verbal warnings as the origin of their fear. It is important to note that those with the conditioning origin rated themselves more fearful on a specific fear questionnaire for their particular animal, and they tended to be more plagued by physiological than by cognitive symptoms. As regards social phobics, over 50% of these reported traumatic incidence as the start of their phobia, with one eighth reporting vicarious experiences, and about 25% being unable to recall how the phobia started. Social phobics with a conditioning history were more plagued by cognitive than by physiological symptoms.

Type of phobia may be related to differential remission of physiological and cognitive symptoms, and therapeutic outcome may fail to indicate the differential effects on the systems. In addition, of course, there are many other factors that determine the outcome of therapy, and these would have to be equated for the results to be regarded as completely negative. It is, after all, the object of a laboratory investigation to eliminate extraneous factors, and concentrate attention on the essential features of the theory; this is in the nature of things impossible in clinical studies.

The Uppsala paradigm may perhaps be extended to the topography of avoidance responses, that is, the proposal that avoidance behavior is subject to adaptive-evolutionary constraints just as much as is the acquisition of conditioned responses. There exists an equipotentiality postulate in the field of avoidance responses as well as in that of conditioned stimuli, stating that any response can serve an avoidance response about as equally as any other response. However, the work of Bolles (1970, 1971, 1972) has shown that response topography may be a critical factor in avoidance behavior, suggesting that the topography of the avoidance response must be in congruence with the behaviors that permitted that organism's ancestors to survive threats in their natural environment. He also assumed that the rate of avoidance learning is determined by how closely the topographical response requirements correspond to these inherited defence repertoires rather than by the functional relationships between incident events, behaviors, and behavioral consequences in the situation.

Seligman and Hager (1972) pointed out an elementary problem with the kind of research on which Bolles has based his views, namely that it did not rule out the type of circular reasoning that is often connected with adaptive-evolutionary hypotheses. The problem arises because criteria for species-specific defense reactions that are independent of avoidance responding need to be specified to prevent experimenters from using learning to define species-specific defense reactions. This and other difficulties led Bolles (1975) to introduce a cognitive element into the adaptive-evolutionary hypothesis. He postulated that responses are not directly evoked by cues predictive of danger but that the subject first learned "expectancies" that certain cues predict shock and that other cues predict safety. Once the animal has been conditioned to expect shock or safety in response to these cues, defensive behavior is automatically released and, consequently, no response learning is required. However, as Delprato (1980) pointed out

this cognitive account compounds the circularity problem noted above in the noncognitive theory. Now if an organism fails to learn to avoid with a particular response x we can maintain either that the response is not a species-specific defence reaction or that the proper expectancies were not learned. If it does learn to avoid with response x , we are free to cite x as a species-specific defence reaction that was evoked by the learned expectancies regarding danger and safety. The circularity involving the hypothetical expectancies is virtually impossible to penetrate because we can never specify an animal's expectancy independently of its behavior. (p. 92)

In summary, it is always possible to argue that when what are regarded as innate defensive reactions, which can be learned easily as avoidance responses, have been contrasted with responses less well learned in the past, that the relevant point is not innateness, but frequency of prior learning. The influence of subtle training conditions on avoidance behavior has been clearly shown by Ayers, Benedict, Glackenmeyer, and Mathews (1974), who controlled training variables that were

confounded with response requirements (head-poke versus lever-press), and found no variance to be left that required genetic considerations.

Altogether, it seems reasonable to accept Delprato's (1980) cautious conclusion that

problems of circular reasoning, failure to control subjects' developmental histories, efficient lever-press avoidance produced in recent experiments, failure to consider training conditions, and confounding of environmental variables with response topography suggest limitations on the adaptive-evolutionary of avoidance learning. (p. 93)

In spite of this necessary caution, the balance of probabilities still indicates a strong genetic effect on response topography.

In his very persuasive critical review of the concept of hereditary determinance of fears and phobias, Delprato (1980) argued against the hypothesis of innate fears and phobias. One phenomenon that has frequently been cited in this connection (e.g., Gray, 1971) is the "hawk-goose" effect, in which a model designed to look like a hawk (short-neck end) when flown in one direction and like a goose (long-neck end) when flown in the opposite direction, is shown to elicit escape reactions, such as flight or crouching in turkey chicks, only when the model is moved in the hawk (predator) direction; the chicks display "merely superficial interest" when the model is moved in the goose (nonpredator) direction (Tinbergen, 1948). Unfortunately, later experiments introducing better control over rearing conditions have given negative results (Hirsch, Lindley, & Tolman, 1955; Rockett, 1955; Schaller & Emlen, 1962). It would seem, rather, that the type of movement (swooping), speed of movement, and stimulus size are more reliably related to fear than are qualitative stimulus characteristics, such as the hawk-goose contrast (Melzack, Penick, & Beckett, 1959; McNiven, 1960; Schleidt, 1961).

Delprato develops his argument in the direction of rejecting theories of innate fear responses, while yet agreeing that conditioning experiences are not sufficient to account for all, or even the majority of fear responses. It seems clear that Delprato is arguing against a conception of behavioral genetics that is not held by behavioral geneticists, but is merely a product of the imagination of psychologists ignorant of the genetic literature. He quotes a long list of critics, mostly going back to the days when behavioral genetics hardly existed as a separate discipline, and still uses these authors as if they had anything relevant to say to present issues. It should be clear that no modern geneticist would claim that there are really innate fears that are not influenced in any way by the developmental history of the organism. It is surely time such outdated views and criticisms were abandoned by anyone concerned with the true genesis of fear (or any other kind of behavior). At least Delprato would seem to agree that the pristine notion of a *tabula rasa* is completely unacceptable, and that there are strong genetic predispositions that must be taken into account by anyone interested in the genesis of human or animal behavior.

CONCLUSIONS

It will be clear from what has been said so far that modern research demands considerable modifications in the theory proposed by Watson and his followers as far as the origin and treatment of neurosis is concerned. It might have been meaningful

in his day and age to dismiss genetic factors, disregard individual differences, and treat the brain as a black box, the structure and functioning of which could tell us nothing useful about behavior, learning, and extinction. This certainly is true no longer, and the fact that ideas of this kind are still common among behaviorists (see Zuriff, 1985, for an enlightened discussion) simply demonstrates that some behaviorists are still ignorant of modern developments.

Equally to be rejected, however, are simple-minded notions of phobic fears, anxieties, etc., being inherited in a direct fashion, uninfluenced by environmental influences, and fixed in a deterministic manner. Genetic factors work in a much more subtle manner, predisposing the organism to react in certain ways to certain environmental stimuli, but certainly not constraining the organism to behave in a preordained manner. The model we have to work with is the threshold model, outlined in Figure 2. The ordinate *neuroticism* indicates differential degrees of predisposition toward the development of a fully blown neurosis; the distribution of this trait is indicated by the normal probability curve over the ordinate. The cross-hatched area indicates cases of actual neurosis, and the stippled line *P* indicates the increasing probability of a person developing a neurosis as his score on neuroticism is further to the right. There is no predetermination in this, only differential probabilities for different people to develop neurotic disorders under environmental stress.

This genetic model emphatically rejects the proposition that environmental factors are unimportant; no meaningful genetic model of human or animal behavior ever made such an assumption. The extremely subtle integration of genetic and environmental factors requires detailed analysis in each case, for each population,

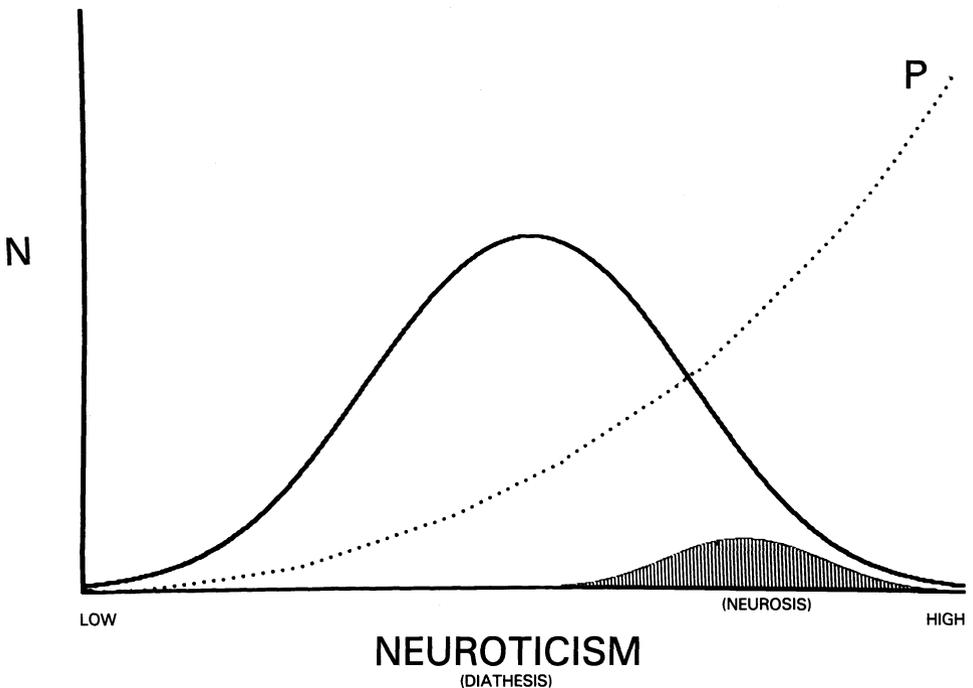


FIGURE 2. The "threshold model" of genetic determination.

and for each trait or ability; simple statements of heritability are only the first step in the solution of a very complex problem that requires consideration of many factors on the genetic and the environmental sides.

In a sense the distinction between inherited fears and prepared fears is unrealistic. There are no fears that are completely inherited; genetic influences can only prepare the organism for the speedy conditioning or learning of specific fear stimuli and fear responses, so that this preparedness constitutes the ordinate in another, not unrelated threshold model. It is curious that Seligman and the Uppsala School, in advocating their model of preparedness, fail to mention the most significant evidence in its favor, namely the genetic evidence.

We have in this chapter been more concerned with the origins rather than with the extinction of neurotic fear responses, and it must be admitted that little direct research has been done on the relationship between genetic and prepared factors and the extinction of fear responses, other than the work of the Uppsala School. It has always been the writer's view that one of the major defects of modern behavior therapy has been the neglect of individual differences and personality. An example, taken from a rather unusual field of application of behavior therapy, may illustrate this point. There is good evidence that personality factors are correlated with cancer and cardiovascular disease, and seem to act in a predisposing fashion, along genetic lines (Eysenck, 1985b). Repression of emotion and learned helplessness seem to be among the most widely noticed of such predisposing factors. The personality of the cardiovascular-disease-prone person is different from that of the cancer-prone person, and in some ways its opposite, containing elements of anxiety, anger, and aggressiveness. Four types were constructed on the basis of observation and descriptions in the literature, Type 1 being theoretically cancer prone, Type 2 prone to cardiovascular disease (CHD), and Types 3 and 4 being relatively healthy and not prone to either cancer or coronary heart disease (Grossarth-Maticek, Eysenck, Vetter, & Frentzel-Beyme, 1986). Each person's type was determined at the beginning of three long-term follow-up studies, and causes of death were ascertained after a 10-year follow-up period. The first of these studies was carried out in Yugoslavia, selecting the oldest person in every second household in a small village. The second study was carried out in Heidelberg on a normal (unselected) group, and the third study was carried out in Heidelberg on a "stress group" nominated by the members of the Heidelberg normal group.

The results of these follow-up studies are shown in Figures 3, 4, and 5 (Eysenck, 1987). It will be seen that as expected people of Type 1 die far more frequently of cancer than members of any of the other groups, whereas people of Type 2 die far more frequently of coronary heart disease. Those of Types 3 and 4 have a very low mortality rate. The figures in the Yugoslav and the Heidelberg stressed-group study are very similar; those in the Heidelberg normal-group study show a much lower degree of disease, which is understandable because of the younger age and lower stress experienced by this group.

Devising methods of behavior therapy to alter the behavior characteristic of personality Types 1 and 2, in a direction opposite to that manifested, Grossarth-Maticek, Schmidt, Vetter, and Arndt (1984) and Grossarth-Maticek, Eysenck, Vetter & Frentzel-Beyme, (1986) showed that it was possible prophylactically largely to

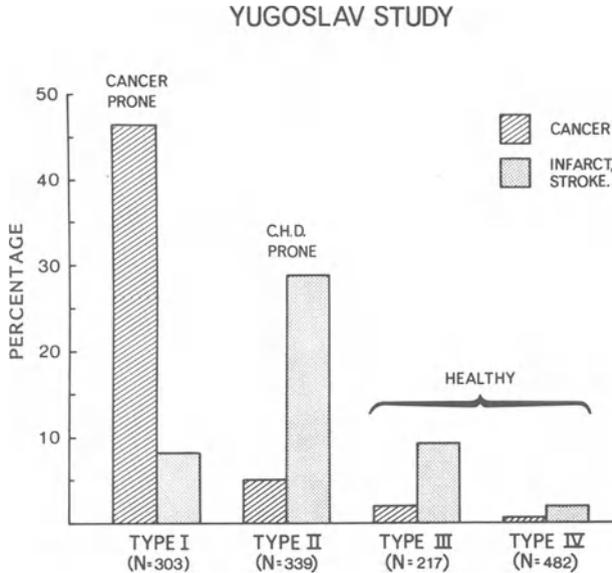


FIGURE 3. Deaths from cancer and coronary heart disease of different personality types. A Yugoslav study.

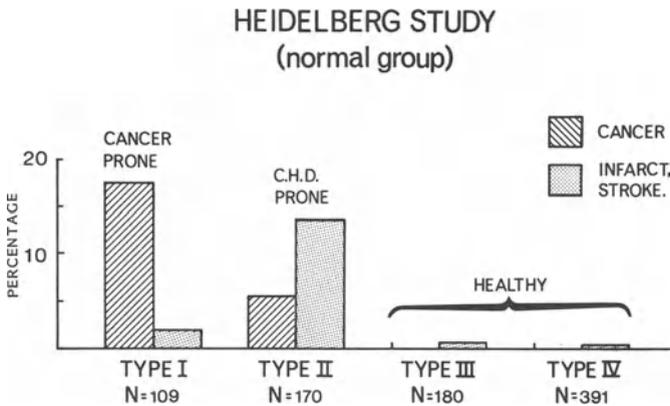


FIGURE 4. Deaths from cancer and coronary heart disease of different personality types. Heidelberg normal-group study.

prevent death from cancer and cardiovascular disease in predisposed groups (when compared with nontreated control groups), thus demonstrating the importance of personality factors even in physical disease and its treatment. They were also successful in prolonging life, using the same methods, in terminally ill cancer patients; their success was equal to that of chemotherapy, and a combination of chemotherapy and behavior therapy had a synergistic effect in prolonging life. Eysenck (1985b, in press) has suggested causal mechanisms to explain just how behavior therapy mediates these effects.

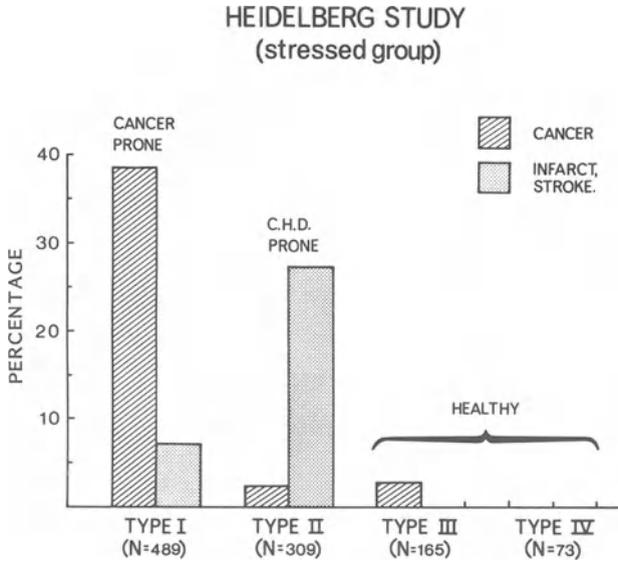


FIGURE 5. Deaths from cancer and coronary heart disease of different personality types. Heidelberg stressed-group study.

The purpose of this brief introduction of a field of work where behavior therapy had not previously been employed is to illustrate the importance of personality factors in devising specific methods of treatment of specific target groups. Traditional behavior therapy, also tried on these groups, was not particularly successful, and certainly much less so than the special types of behavior therapy based on the analysis of genetic personality types. It seems likely that if behavior therapists were to pay more attention to personality and individual differences in the treatment of neurotic disorders, they might be more successful than they are at present.

REFERENCES

- Agras, W. S., Sylvester, D., & Olivean, D. (1969). The epidemiology of common fears and phobias. *Comprehensive Psychiatry*, *10*, 151-156.
- Arrindell, W. A. (1980). Dimensional structure and psychopathology correlates of the Fear Survey Schedule (FSS-111) in a phobic population: A factorial definition of agoraphobia. *Behaviour Research and Therapy*, *18*, 229-242.
- Arrindell, W. A., & van der Ende, J. (1986). Further evidence for cross-sample invariance of phobic factors: Psychiatric inpatient ratings on the Fear Survey Schedule - 111. *Behaviour Research and Therapy*, *24*, 289-297.
- Arrindell, V. A., & Zwaan, E. J. (1985). Fobische dimensies de AZV-Vreesvragenlijst (AZUV) en Wolpen Lang's Fear Survey Schedule (FSS-111). *Ned. Tijdsch. Psychol.*, *9*, 165-178.
- Arrindell, A., Emmelkamp, P. M. G., & van der Ende, J. (1984). Phobic dimensions: I. Reliability and generalisability across samples, gender and nations. *Advances in Behaviour Research and Therapy*, *6*, 207-253.

- Ayers, J., Benedict, J., Glackenmeyer, R., & Mathews, W. (1974). Some factors involved in the comparison of response systems: Acquisition, extinction and the transfer of head-poke and lever-press Sidman avoidance. *Journal of the Experimental Analysis of Behaviour*, 22, 371-379.
- Bedford, A., Foulds, G. A., & Sheffield, B. F. (1976). A new personal disturbance scale (DSSI/sAB). *British Journal of Social and Clinical Psychology*, 15, 387-394.
- Bolles, R. C. (1970). Species-specific defense reactions and avoidance learning. *Psychological Review*, 77, 32-48.
- Bolles, R. C. (1971). Species-specific defense reactions. In F. R. Brush (Ed.), *Aversive conditioning and learning*. New York: Academic Press.
- Bolles, R. C. (1972). The avoidance learning problem. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 6, New York: Academic Press.
- Bolles, R. C. (1975). *Theory of motivation* (2nd ed.). New York: Harper & Row.
- Cabe, P. A. (1976). Transfer of discrimination from solid objects to pictures by pigeons: A test of theoretical models of pictorial perception. *Perception and Psychophysics*, 19, 545-550.
- Carey, G., & Gottesman, I. I. (1981). Twin and family studies of anxiety, phobic, and obsessive disorders. In D. F. Klein & J. Rabkin (Eds.), *Anxiety: New research and changing concepts* (pp. 117-136). New York: Raven Press.
- Cerella, J. (1979). Visual classes and natural categories in the pigeon. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 68-77.
- Comings, D. E. (1979). P_u1 Duarte, a common polymorphism of a human brain protein, and its relationship to depressive disease and multiple sclerosis. *Nature*, 277, 28-32.
- Cook, E. W. (1981). *Classical conditioning, individual differences, and preparedness*. Unpublished master's thesis, University of Wisconsin, Madison, WI.
- Cook, E. W. (1983). *Human classical conditioning and the preparedness hypothesis*. Unpublished doctoral dissertation, University of Wisconsin, Madison, WI.
- Dawson, M. E., & Schell, A. M. (1985). Electrodermal classical conditioning with potentially phobic CSs: Cognitive autonomic dissociations?
- Dawson, M. E., & Schell, A. M. (1985). Information processing and human autonomic classical conditioning. In P. K. Ackles, J. E. Jennings & M. G. Coles (Eds.), *Advances in psychophysiology*, Greenwich, CT: JAI Press.
- Delprato, D. J. (1980). Hereditary determinants of fears and phobias: critical review. *Behaviour Therapy*, 11, 79-103.
- De Silva, P., Rachman, S., & Seligman, M. E. P. (1977). Prepared phobics and obsessions: Therapeutic outcome. *Behaviour Research and Therapy*, 15, 65-77.
- Dixon, J. J., de Monchaux, C., & Sandler, J. (1957). Patterns of anxiety: The phobias. *British Journal of Medical Psychology*, 30, 34-40.
- Doctor, R. M. (1982). Major results of a large-scale pretreatment survey of agoraphobics. In R. L. DuPont (Ed.), *Phobia: A comprehensive summary of modern treatments*. New York: Bruner-Mazel.
- Eaves, L. J., & Eysenck, H. J. (1975). The nature of extraversion: A genetical analysis. *Journal of Personality and Social Psychology*, 32, 102-112.
- Eaves, L. J., & Eysenck, H. J. (1976a). Genotype \times age interaction for neuroticism. *Behaviour Genetics*, 6, 359-362.
- Eaves, L. J., & Eysenck, H. J. (1976b). Genetical and environmental components of inconsistency and incompatibility in twins' responses to a neuroticism questionnaire. *Behaviour Genetics*, 6, 145-160.
- Eaves, L. J., & Eysenck, H. J. (1977). A genotype environmental model for psychoticism. *Advances in Behaviour Research and Therapy*, 1, 5-26.
- Eaves, L. J., & Eysenck, H. J. (in press). *The genetics of personality*. London: Academic Press.
- Eaves, L. J., & Young, P. A. (1981). Genetical theory and personality differences. In R. Lynn (Ed.), *Dimensions of personality* (pp. 129-179). London: Pergamon Press.
- Eysenck, H. J. (1947). *Dimensions of personality*. London: Routledge & Kegan Paul.
- Eysenck, H. J. (1952). *The scientific study of personality*. London: Routledge & Kegan Paul.
- Eysenck, H. J. (1956). The inheritance of extraversion-neuroticism. *Acta Psychologica*, 12, 95-110.
- Eysenck, H. J. (1967). *The biological basis of personality*. Springfield, IL: Charles C Thomas.
- Eysenck, H. J. (Ed.). (1981). *A model for personality*. New York: Springer.

- Eysenck, H. J. (1983). A biometrical-genetical analysis of impulsivity and sensation seeking behaviour. In M. Zuckerman (Ed.), *Biological Bases of Sensation Seeking, Impulsivity and Anxiety* (pp. 1–27). Hillsdale, NJ: Erlbaum.
- Eysenck, H. J. (1985a). Personality, cancer and cardiovascular disease: A causal analysis. *Personality and Individual Differences*, 6, 535–556.
- Eysenck, H. J. (1985b). Smoking and health. In R. D. Tollison (Ed.), *Smoking and society* (pp. 17–88). Toronto: Lexington Books.
- Eysenck, H. J. (1987). Anxiety, “learned helplessness” and cancer—A causal theory. *Journal of Anxiety Disorders*, 1, 87–104.
- Eysenck, H. J., & Eysenck, S. B. G. (1975). *The manual of the Eysenck Personality Questionnaire*. London: Hodder & Stoughton, San Diego: EDITS.
- Eysenck, H. J., & Prell, D. (1951). The inheritance of neuroticism: An experimental study. *Journal of Mental Science*, 97, 441–465.
- Floderus-Myrhed, B., Pederson, N., & Rasmuson, I. (1980). Assessment of heritability for personality based on a short form of the Eysenck Personality Inventory. A study of 12898 twin pairs. *Behaviour Genetics*, 10, 153–162.
- Foulds, G. A. (1965). *Personality and personal illness*. London: Tavistock.
- Foulds, G. A. (1974). *The hierarchical nature of personal illness*. London: Academic Press.
- Fulker, D. W. (1981). The genetic and environmental architecture of psychoticism, extraversion and neuroticism. In H. J. Eysenck (Ed.), *A model for personality* (pp. 88–122). New York: Springer.
- Fulker, J. L., & Simmel, E. C. (Eds.). (1983). *Behaviour genetics*. London: Erlbaum.
- Gray, J. (1970). The psychological basis of introversion-extraversion. *Behaviour Research and Therapy*, 8, 249–266.
- Gray, J. (1971). *The psychology of fear and stress*. New York: McGraw-Hill.
- Grossarth-Maticek, R., Schmidt, P., Vetter, H., & Arndt, S. (1984). Psychotherapy research in oncology. In A. Steptoe & A. M. Mathews (Eds.), *Health care and human behaviour* (pp. 325–340). London: Academic Press.
- Grossarth-Maticek, R., Eysenck, H. J., Vetter, H., & Frentzel-Beyme, R. (1986, March). *The Heidelberg prospective intervention study*. Paper presented at the First International Symposium of Primary Prevention of Cancer, Antwerp, Belgium.
- Hafner, R. J., & Ross, M. W. (1984). Agoraphobia in women: Factor analysis of symptoms and personality correlates of factor scores in a clinical population. *Behaviour Research and Therapy*, 22, 441–444.
- Hallam, R. S., & Hafner, R. J. (1978). Fear of phobic patients: factor analyses of self-report data. *Behaviour Research and Therapy*, 16, 1–6.
- Hamm, A., & Vaitl, D. (1985). “Belongingness”: Operationalisierung eines neuen Konzepts in der Lerntheorie. *Archiv für Psychologie*, 137, 13–27.
- Heath, S. C., Berg, K., Eaves, L., Solach, M. H., Currey, L., Sundet, J., Magnus, P., & Vance, W. E. (1985). Education policies and the verifiability of educational attainment. *Nature*, 315, 734–736.
- Herrnstein, R. J. (1979). Acquisition, generalization, and discrimination reversal of a natural concept. *Journal of Experimental Psychology: Animal Behavior Processes*, 5, 116–129.
- Herrnstein, R. J. (1980). Fish as a natural category for people and pigeons. *The Psychology of Learning and Motivation*, 14, 59–95.
- Herrnstein, R. J. (1984). Objects, categories, and discriminative stimuli. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal Cognition* (pp. 233–261). London: Erlbaum.
- Herrnstein, R. J. (1985). Riddles of natural categorization. *Philosophical transactions of the Royal Society*. London. B 308, 129–144.
- Herrnstein, R. J., & Loveland, D. H. (1964). Complex visual concepts in the pigeon. *Science*, 146, 549–551.
- Herrnstein, R. J., Loveland, D. H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 2, 285–311.
- Hersen, M. (1983). Self-assessment of fear. *Behaviour Therapy*, 4, 241–257.
- Hirsch, J., Lindley, R., & Tolman, E. C. (1955). Experimental test of an alleged innate sign stimulus. *Journal of Comparative and Physiological Psychology*, 48, 278–280.
- Hodes, R. L. (1981). *A psychophysiological investigation of the classical conditioning model of fears and phobias*. Unpublished doctoral dissertation, University of Wisconsin, Madison, WI.
- Hugdahl, K. (1978). Electrodermal conditioning to potentially phobic stimuli. Effects of instructed extinction. *Behaviour Research and Therapy*, 16, 315–321.

- Hugdahl, K., & Kärker, A. C. (1981). Biological vs. experiential factors in phobic conditioning. *Behaviour Research and Therapy*, *19*, 109–115.
- Hugdahl, K., & Öhman, A. (1977). Effects of instruction on acquisition of extinction of electrodermal responses to fear relevant stimuli. *Journal of Learning and Memory*, *3*, 608–618.
- Hugdahl, K., & Öhman, A. (1980). Skin conductance conditioning to potentially phobic stimuli as a function of interstimulus interval and delay versus trace paradigm. *Psychophysiology*, *17*, 348–355.
- Hugdahl, K., Fredrikson, M., & Öhman, A. (1977). "Preparedness" and "arousability" as determinants of electrodermal conditioning. *Behaviour Research and Therapy*, *15*, 345–353.
- Jardine, R., Martin, N. G., & Henderson, A. S. (1984). Genetic covariation between neuroticism and the symptoms of anxiety and depression. *Genetic Epidemiology*, *1*, 89–107.
- Katz, R., & McGuffin, P. (in press). Neuroticism in familial depression. *Psychological Medicine*.
- Kendler, K. S., Heath, S., Martin, N. G., & Eaves, L. J. (1986). Symptoms of anxiety and depression in a volunteer twin population. *Archives of General Psychiatry*, *43*, 213–221.
- Loehlin, J. C., & Nichols, R. C. (1976). *Heredity, environment and personality: A study of 850 sets of twins*. Austin, TX: University of Texas Press.
- Martin, N., & Jardine, R. (1986). Eysenck's contribution to behaviour genetics. In S. Modgil & C. Modgil (Eds.), *Hans Eysenck: Consensus and Controversy* (pp. 13–47).
- Martin, N. G., Eaves, L. J., Kearsley, M. J., & Davies, P. (1978). The power of the classical twin study. *Heredity*, *40*, 97–116.
- Mather, K., & Jinks, J. L. (1971). *Biometrical genetics* (2nd ed.). London: Chapman & Hall.
- McNally, R. D., & Foa, G. B. (1986). Preparedness and resistance to extinction to fear-relevant stimuli: A failure to replicate. *Behaviour Research and Therapy*, *24*, 529–535.
- McNiven, S. A. (1960). "Social releaser mechanism" in birds—A controlled replication of Tinbergen's study. *Psychological Record*, *10*, 259–265.
- Melzack, R., Penick, E., & Beckett, A. (1959). The problem of "innate fear" of the hawk shape: An experimental study of Mallard ducks. *Journal of Comparative and Physiological Psychology*, *52*, 694–698.
- Miner, G. D. (1973). The evidence for genetic components in the neuroses. *Archives of General Psychiatry*, *29*, 111–118.
- Newman, H. H., Freeman, F. N., & Holzinger, K. J. (1937). *Twins*. Chicago, IL: University of Chicago Press.
- Öhman, A. (1979). Fear relevance, autonomic conditioning, and phobias: A laboratory model. In P. O. Sjöden, S. Bates, & W. S. Dockens (Eds.), *Trends in behavior therapy*. New York: Academic Press.
- Öhman, A., & Hugdahl, K. (1979). Instructional control of autonomic respondents: Fear-relevance as a critical factor. In N. Birbaumer & H. D. Kimmel (Eds.), *Bio-feedback and self-regulation*. Hillsdale, NJ: Erlbaum.
- Öhman, A., Eriksson, G., & Löfberg, I. (1975). Phobias and preparedness: Phobic versus neutral pictures as conditioned stimuli for human autonomic responses. *Journal of Abnormal Psychology*, *84*, 11–45.
- Öhman, A., Eriksson, A., & Olöfsson, C. (1975). One-trial learning and superior resistance to extinction of autonomic responses conditioned to potentially phobic stimulus. *Journal of Comparative and Physiological Psychology*, *88*, 619–627.
- Öhman, A., Fredrikson, M., Hugdahl, K., & Rimmö, P. (1976). The premise of equipotentiality in human classical conditioning: Conditioned electrodermal responses to potentially phobic stimuli. *Journal of Experimental Psychology: General*, *105*, 313–337.
- Öhman, S., Dimberg, V., & Öst, L.-G. (1985). Animal and social phobias: Biological constraints on learned fear responses. In S. Reiss & R. R. Bootzin (Eds.), *Theoretical issues in behavior therapy* (pp. 123–175).
- Öst, L. G., & Hugdahl, K. (1981). Acquisition of phobias and anxiety response patterns in clinical patients. *Behaviour Research and Therapy*, *19*, 439–447.
- Parker, G. (1979). Parental characteristics in relation to depressive disorders. *British Journal of Psychiatry*, *194*, 138–147.
- Parker, G. (1981a). Parental reports of depressives: An investigation of several explanations. *Journal of Affective Disorders*, *3*, 131–140.
- Parker, G. (1981b). Parental representations of patients with anxiety neurosis. *Acta-Psychiatrica Scandinavica*, *63*, 33–36.
- Quine, W. V. O. (1969). Natural kinds. In W. V. O. Quine (Ed.), *Ontological relativity and other essays*. New York: Columbia University Press.

- Rachman, S. (1977). The conditioning theory of fear-acquisition: A critical examination. *Behaviour Research and Therapy*, 15, 375–387.
- Rachman, S., & Seligman, M. E. P. (1976). Unprepared phobias: "Be Prepared." *Behaviour Research and Therapy*, 14, 333–338.
- Redlich, F. C., & Freedman, D. (1966). *The theory and practice of psychiatry*. New York: Basic Books.
- Rockett, F. C. (1955). A note on "an experimental test of an alleged innate sign stimulus" by Hirsch, Lindley and Tolman. *Perceptual and Motor Skills*, 5, 155–156.
- Roubertoux, P., & Carlier, M. (1972, 1973). Analyse génétique des comportements. *L'Année Psychologique*, 72, 393–441; 73, 151–223.
- Rushton, J. P., Fulker, D. W., Neale, M. C., Nias, D., & Eysenck, H. J. (1986). Altruism and aggression: The heritability of individual differences. *Journal of Personality and Social Psychology*, 50, 1192–1198.
- Schaller, G. B., & Emlen, J. T. (1962). The autogeny of avoidance behaviour in some precocial birds. *Animal Behaviour*, 10, 370–381.
- Schepank, K. H. (1973). Erb-und Umweltfaktoren bei Neurosen. *Nervenarzt*, 44, 449–459.
- Schleidt, W. M. (1961). Reaktionen von Truthütern auf fliegende Raubvögel und Versuche zur Analyse ihrer AAMs. *Zeitschrift für Tierpsychologie*, 18, 534–560.
- Seligman, M. E. P. (1971). Phobias and preparedness. *Behavior Therapy*, 2, 307–321.
- Seligman, M. E. P., & Hager, J. C. (1972). *Biological boundaries of learning*. New York: Appleton-Century-Crofts.
- Shields, J. (1962). *Monozygotic twins*. Oxford University Press.
- Slater, E. (1943). The neurotic constitution: A statistical study of two thousand neurotic soldiers. *Journal of Neurology, Neurosurgery and Psychiatry*, 6, 1–16.
- Slater, E., & Shields, J. (1969). Genetical aspects of anxiety. In M. Lader (Ed.), *Studies of anxiety*. Ashford, Kent: Headley.
- Slater, E., & Slater, P. (1944). A heuristic theory of neurosis. *Journal of Neurology, Neurosurgery and Psychiatry*, 7, 49–55.
- Tasto, D. L. (1977). Self-report schedules and inventories. In A. R. Ciminero, K. S. Calhom, & H. E. Adams (Eds.), *Handbook of behavioral assessment*. New York: Wiley.
- Thorndike, E. L. (1931). *Human learning*. New York: D. Appleton-Century.
- Tinbergen, N. (1948). Social releases and the experimental method required for their study. *The Wilson Bulletin*, 60, 6–51.
- Torgersen, S. (1979). The nature and origin of common phobic fears. *British Journal of Psychiatry*, 134, 343–351.
- Vandenberg, S. G. (1969). Hereditary factors in normal personality traits (as measured by inventories). *Recent Advances in Biological Psychiatry*, 9, 65–104.
- Wade, T. C. (1978). Factor analytic approaches to the investigation of common fears: A critical appraisal and reanalysis. *Behavior Therapy*, 9, 923–935.
- Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology*, 3, 1–14.
- Weitkamp, L. R., Stanger, H. C., Persad, E., Flood, C., & Guttormsen, S. (1981). Depressive disorder and HLA: A gene on chromosome 6 that can affect behaviour. *New England Journal of Medicine*, 305, 1301–1306.
- Wolpe, J., & Lange, P. J. (1964). A fear survey schedule for use in behavior therapy. *Behavior Research and Therapy*, 2, 27–30.
- Zuriff, G. E. (1985). *Behaviorism: A conceptual reconstruction*. New York: Columbia University Press.

CHAPTER 19

Hormones and Clinical Anxiety

An Imbalanced Neuromodulation of Attention

Michael J. Kelley

INTRODUCTION

In this chapter, a model of neuroses will be developed that suggests that clinical anxiety is an outcome of an interaction between associative conditioning and individual differences in the functioning of the neuroendocrine system. The model takes as an assumption that there is a synergism between these two factors and that each alone is largely insufficient to produce clinical neuroses. More broadly, this interaction can be conveyed by the following equation:

$$\text{neuroses} = \text{conditioning} \times \text{neurohormones}$$

This formulation can be contrasted with models of neurotic behavior that suggest conditioning factors alone are sufficient for the development of neurotic behavior. Although there is ample evidence that the occurrence of stressful events plays an important role in the development of neuroses, most people recover from such experiences without showing protracted neurotic behavior of a clinical magnitude. A traditional way out of this dilemma is to propose that some individuals have characteristics that make them more susceptible to neuroses after such conditioning. In the present formulation, individual differences in the neuroendocrine system serve this function. It will be argued that aversive conditioning in conjunction with these predispositions results in anxiety that is highly persistent and that can become insidiously worse following mere exposure to stimuli associated with the trauma. These are the two key characteristics of neurotic anxiety that conditioning theories of neuroses have struggled to explain (Eysenck, 1979).

This model also generates testable predictions for improved treatment of clinical neuroses. Exposure to aversive cues in conjunction with abnormal levels of hormones need not always lead to a permanent increase in the excitatory properties of these stimuli. Most hormones are probably irrelevant to neuroses, but alterations in the

levels of certain hormones may produce circumstances that allow a fear cue to rapidly lose its capacity to elicit anxiety. In this regard, much of the future evidence to support this model will probably come from studies showing that more efficacious treatment of neurotic disorders occurs when neurohormonal manipulations are part of the therapeutic conditioning procedures known as behavior therapy. This again can be phrased as an equation: behavior therapy = conditioning \times neurohormones. As in the earlier equation, the multiplication symbol is the key element. It suggests that the impact of therapeutic conditioning procedures, such as the protracted exposure to phobic stimuli, may be substantially enhanced by concurrent manipulation of hormone factors, and possibly more important, that the manipulation of only one factor may be insufficient or largely ineffectual.

The study of interactions between hormones and anxiety has a long history, thus it is important to delineate where the present model differs from previous ones. Following Cannon's (1915) demonstration that epinephrine is secreted in response to stress, Tompkins, Sturgis, and Wearn (1919) showed that neurotics were more likely than normal individuals to respond with an increase in anxiety after an epinephrine injection. Subsequently, in a study with normal individuals, Schacter and Singer (1962) showed that epinephrine itself was not anxiogenic and that any such effect was dependent on the emotional properties of the prevailing stimuli. This situation appears also to be true for many other neurohormones that have been subsequently discovered to influence anxiety, that is, only the hormone plus a fear stimulus produces greater anxiety. The present model goes one step further and assumes that contiguity between the presence of anxiety cues and an imbalance in the neuroendocrine system can produce a permanent increase in the capacity of these cues to elicit anxiety. These cues may be internal sensations produced by anxiety itself, as suggested by Breggin (1964), and also certain potentially phobic external cues that are now known to form rapidly associations with fear. This process by which fear cues can acquire a permanent increase in their capacity to elicit responses has been called the incubation of anxiety by Eysenck (1979). The present model merely specifies some hormonal factors that appear to be permissive conditions for incubation to occur.

The chapter begins with an introduction to neurohormones and some basic psychological concepts. Following this the animal evidence is reviewed, which suggests that there is a balanced hormonal modulation of extinction and incubation, and that this modulation occurs by effects on attention. The empirical work with humans is then examined, which implicates the actions of hormones on attention and supports the proposition that hormones may influence the developmental course of neurotic behavior.

SOME BASIC CONCEPTS

Although the terms *conditioning*, *hormones*, and *neuroses* are familiar concepts to the readers of this chapter, the diversity of definitions in the scientific literature, to say nothing of the popular literature, demands that some clarification be first given to these terms.

DEFINITIONS OF CONDITIONING, ATTENTION, AND NEUROSES

A conditioned stimulus (CS+) with excitatory properties for fear is formed when the cue is paired with an unconditioned stimulus (UCS). The neurochemical changes elicited by the UCS is what causes the organism to change its responsiveness to the CS+ (McGaugh, 1983); however, an aversive UCS can also be defined by its capacity to elicit defensive behavior. In humans (and many other mammals), an effective UCS is often the defensive reactions of another person to a CS+, hence the previous associative definition of conditioning also encompasses what is called observational learning. In what follows, the terminology of Pavlovian conditioning will be used, even when the investigators describe their experiments in what is called instrumental conditioning procedures. The justification is simple; although the procedural terminology may differ, the underlying associative processes are for the most part, if not entirely, the same in both instrumental and Pavlovian conditioning (Mackintosh, 1983). This appears to be true for the effects of hormones on conditioning.

Not all associative learning involves changes in the excitatory strength of a stimulus. If a novel stimulus is repeatedly presented to an animal, it is then difficult to transform this stimulus into an effective CS+. This acquired constraint on learning is known as latent inhibition, and learned irrelevance when the limitation on future conditioning is produced by random presentations of the UCS and the stimulus. Less technically, we can say the subject has learned to ignore the stimulus or response. There are reasons to believe that this same diminution of attention occurs when a CS+ no longer predicts the occurrence of a UCS (Mackintosh, 1983). One of the important ways that neurohormones can influence anxiety is by disrupting or enhancing the normal effects of these procedures.

In the animal psychopharmacological literature that will be discussed shortly, a change in reactivity to an excitatory conditioned stimulus (CS+) is often used as a basis for saying the animal's memory has been influenced, but this by itself is not an adequate explanation, as several different processes may produce this change in performance. The phrase *enhanced extinction* will be used here to describe an experimental outcome in which a CS+ declines in strength faster, and the phrase *reduced extinction* will be used when this occurs more slowly. Another possible outcome sometimes occurs after CS+ is presented alone after prior conditioning; instead of declining in strength the CS+ may show a permanent increase in excitatory strength (the opposite of extinction). This is called incubation. It is important to realize that the terms *enhanced consolidation*, *retention*, and *retrieval* are often used in the psychopharmacology literature when changes in extinction would be a more parsimonious description of the empirical results. Memory, of course, may be involved but this is a theoretical inference about causal processes and this should be separated from the description of the experimental results themselves.

Neuroses are extreme expressions of species-typical defensive reactions. This is to say that neurotic behavior has an identifiable structure that allows for the construction of behavioral inventories in which there are substantial correlations between the items. In this regard, the Neuroticism factor in the Eysenck Personality Questionnaire (EPQ) can be seen as an example of what biologists call an ethogram (i.e., complete behavioral description) of human defensive behavior (Eysenck & Eysenck,

1985). Clinical neuroses are the extreme, protracted changes in the expression of one or more of these items. There are many reasons for emphasizing this biological side of neuroses. For instance, in the analysis that follows a causal link between individual differences in hormones and anxiety is proposed in which there is already a basis for suspecting strong genetic mediation (Zerbe, 1985). Another tie between neuroticism and the species-typical defensive behavior of animals is that the entire defensive repertoire is changed when fear is increased in animals by environmental, neurobiological, or genetic factors (Bolles, 1970; Eysenck & Eysenck, 1985); this is consistent with the common observation that there is great overlap between the many diagnostic categories of anxiety-related disorders, to say nothing of developmental shifts between them. Because of this, the focus here will be to deal with elements that are common to all of them, primarily the persistence and incubation of anxiety.

There is another important tie between species-typical defensive behavior and human anxiety. Conditioning is largely ineffective in altering the form of defensive behavior itself; however, it does change the stimulus circumstances under which defensive behavior is expressed and the magnitude of this expression. The animal evidence clearly shows that species-typical behavior determines what is most readily learned (Grossen & Kelley, 1972; Kelley, 1985b). This constraint on conditioning occurs not only with behavior but also with the stimuli that control these reactions in humans and animals, that is, only a limited number of stimuli readily become associated with aversive events (Kelley, 1986a; Öhman, Dimberg, & Ost, 1985). This constraint on stimulus associability can be best explained in terms of attention theory (Mackintosh, 1983). In the development of clinical anxiety, it is likely that the hormone factors described here operate in conjunction with these constraints on associations.

A DEFINITION OF NEUROHORMONES

Many hormones are known by some function they serve in the body, but to identify one function in one portion of the body does not exclude the possibility of a different function elsewhere. This is a very important lesson to remember when reading even recent texts on physiological psychology or endocrinology, because many of the classical hormones serve other functions in the brain. There is evidence that the hormones oxytocin and vasopressin serve a role as neurotransmitters in the brain (Buijs, 1983). It is not yet known just how many hormones are also neurotransmitters, but it is certainly clear that the six or so traditional neurotransmitters that have formed the corner stone of clinical neuroscience during the last few decades are only part of the story, possibly a very small part (Iversen, 1984; Krieger, 1983). Aside from sometimes also being neurotransmitters, there are other roles for hormones in the brain that may be equally important. For instance, they may influence metabolism, thresholds for neural conduction, and the turnover rates of other neurotransmitters. Evidence of these other CNS functions has given rise to a new term, *neuromodulator*, to refer to the broad class of substances that can alter information flow in the nervous system without themselves necessarily being neurotransmitters.

NEUROHORMONES AND THE EXTINCTION OF ASSOCIATIONS IN ANIMALS

There is extensive evidence from animal learning experiments which shows that levels of neurohormones are a significant factor in the extinction of aversive conditioning (Van Wimersma Greidanus *et al.*, 1983). The fact that many hormones appear to be involved and often with opposing effects has led to the view that there is a balanced endogenous neuromodulation of extinction. An overview of this area will now be provided but no attempt will be made to cover all the possible hormones that may be involved. The focus is on extinction rather than the acquisition of conditioning because this is where the most dramatic effects have been observed.

CORTICOTROPIN AND RELATED PEPTIDES

The bulk of the early work on hormones and extinction has been done with corticotropin, which is also known as adrenocorticotrophic hormone (ACTH). This hormone is called a peptide because it is composed of a sequence of amino acids. The number of amino acids involved in a peptide is less than that of a protein but the distinction between the two is arbitrary.

Enhanced resistance to extinction after peripheral or central administration of ACTH occurs in a variety of aversive conditioning tasks. Although it is also possible to demonstrate the effects of ACTH on the extinction of behavior reinforced with appetitive (i.e., food or water) reinforcers, the results are less dramatic and largely difficult to replicate. Outside of acquired fear, sexual behavior is the only other motivational system where robust effects are observed; this is probably because only these two motivational systems are capable of supporting learning based on drive induction as well as drive reduction (Eysenck & Kelley, 1986). Figure 1(a) shows a recent demonstration of the effect of pre-session ACTH injections on the extinction of shuttle-box avoidance in rats (De Vito & Brush, 1984). A comprehensive review of this literature has been recently published by de Wied and Jolles (1982).

One natural function of ACTH when it is released from the pituitary is to promote the release of cortisol from the adrenal glands; however, the effect of ACTH on extinction is independent of its capacity to cause the release of cortisol. Evidence for this comes from studies showing an effect on extinction even when the adrenal glands are removed and when some synthetic peptide fragments of this hormone (ACTH4-10 and ACTH4-9) that lack adrenal properties are administered. The notation 4-10 and 4-9 means that only these sequences of amino acids are contained in these peptide fragments, whereas the natural hormone, ACTH(1-39), contains a longer sequence. Evidence of a natural physiological role of ACTH in the CNS comes from studies showing that ACTH exists in the CNS and that there is a reduced extinction when ACTH antiserum is injected into limbic structures. This demonstration of a separation of behavioral from endocrine functions has produced a revolution in the psychopharmacology of hormones. Other synthetic fragments of ACTH that are devoid of endocrine properties have the opposite effect on extinction

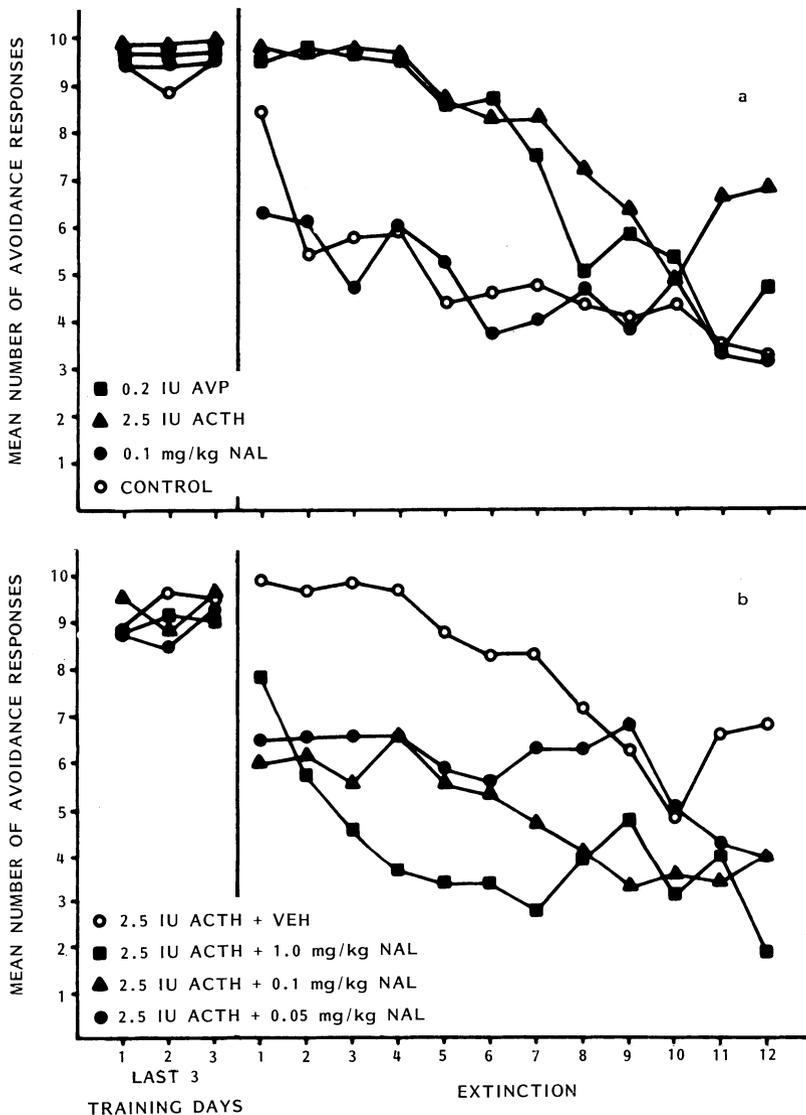
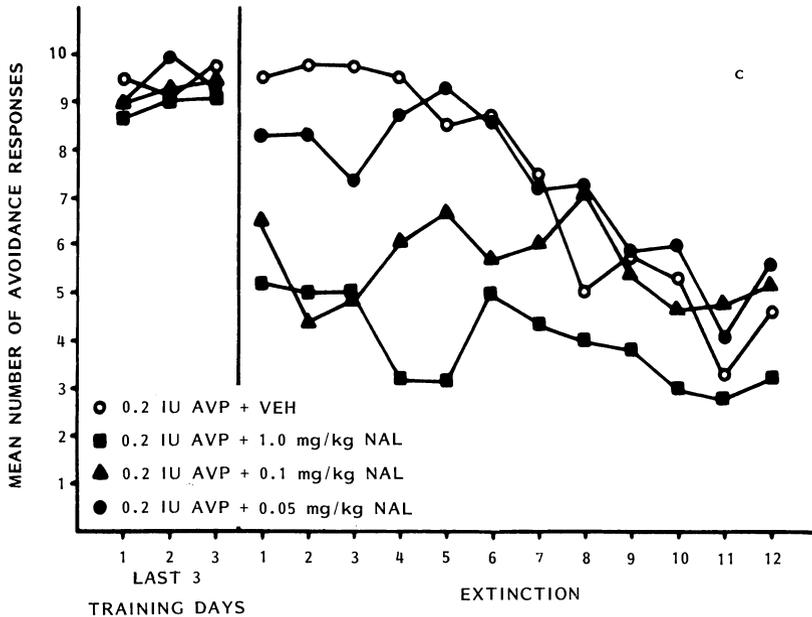


FIGURE 1. The effects of daily pre-session administration of saline, ACTH or VP in the testing for extinction of shuttle-box avoidance responding in rats. The influence of ACTH and VP on reducing extinction is blocked by giving prior injections of naloxone. The figure is reprinted from DeVito and Brush (1984).

but considerably less work has been done with them (Meyer & Bohus, 1983). The behaviorally active core sequence of amino acids in ACTH is also contained in another hormone called alpha melanocyte-stimulating hormone (alpha-MSH). It has much the same influence on extinction as ACTH(4-10), exists as a natural metabolite in the brain, and will be discussed again in the section on neurohormones and neuroses.

FIGURE 1. *Continued.*

CORTISOL AND CORTICOSTERONE

The opposite effect to ACTH, more rapid extinction, occurs after the administration of the adrenal steroids corticosterone or cortisol. This has been explored with a variety of aversive conditioning procedures; but from the standpoint of behavior therapy, one of the most interesting procedures is a three-stage experimental design in which the second stage is analogous to the behavior therapy procedure called flooding. In the first stage an excitatory CS+ is established by pairing it with an aversive UCS. In the second stage the animal is exposed again to the CS+. In the third stage, at least a day later, the impact of the prior nonreinforced exposure is assessed by some behavioral measure of fear. In the absence of neuroendocrine manipulations, the typical outcome of exposure to the CS+ in the second stage is a permanent decrease in fear of the CS+, which is observed during the third stage.

Bohus and his collaborators have drawn attention to the fact that hormonal manipulations can markedly alter the amount of extinction to the CS+ that occurs in the second stage when the animal is exposed to the CS+ (Bohus, De Kloet, & Veldhuis, 1982). They have shown that if the adrenal glands of rats were removed (adrenalectomy) prior to stage two, this resulted in no extinction occurring to the CS+. Although this is strong evidence that a neuroendocrine manipulation can influence the outcome of forced exposure to a CS+, the experiment by itself is ambiguous as to what hormones are involved; adrenalectomy reduces corticosteroid levels and also produces a marked rise in ACTH levels, among other things. Other experiments by the same laboratory showed that if corticosterone was implanted into the limbic system of the adrenalectomized rats, this normalized extinction during

flooding, whereas it did not influence levels of ACTH. This effect can also be produced by peripheral injections of corticosterone. Other steroid hormones, such as progesterone or dexamethasone, do not produce this effect, but they are capable of occupying the corticosterone receptors in the CNS and thus prevent the normalization of extinction by corticosterone.

CORTICOTROPIN-RELEASING FACTOR

The structure of the principle hypothalamic peptide that causes the release of ACTH has been recently identified and is called corticotropin-releasing factor (CRF). Like ACTH its level in plasma and cerebrospinal fluid (CSF) can be determined by radioimmunoassay and it is available in synthetic form along with antagonists to its effects. It is the focus of considerable current research (Gold, Chrousos, *et al.*, 1984; Heumann, 1985). One group of investigators has found that stress increases CSF levels of this peptide, and also repeatedly observed that intracerebroventricular injections suppress appetite and directly potentiate defensive behavior in animals; however, this influence on emotionality has not been replicated by another group (Britton, Varela, Garcia, & Rosenthal, 1986; Veldhuis & de Wied, 1984). In general, direct effects of CRF and ACTH/MSH on emotionality have been difficult to replicate in animals and even more so in humans, but the issue is far from resolved (Beckwith & Sandman, 1978, 1982; Datta & King, 1982; Gold, Chrousos, *et al.*, 1984).

BETA-ENDORPHIN

There is now considerable evidence that beta-endorphin (B-E), one of the principal endogenous opioids, is released in the CNS during exposure to a variety of different kinds of aversive events; in addition, direct administration of physiological amounts of B-E into the CNS produces an attenuation of acquired fear reactions. This is blocked by prior injections of the opiate antagonist naloxone. The existence of cross tolerance between exogenous opiates and the endogenous B-E released during stress is now well established (Fanselow, 1985; Izquierdo *et al.*, 1981; Lester & Fanselow, 1985; Terman, Shavit, Lewis, Cannon, & Liebeskind, 1984; Williams, Drugan, & Maier, 1984). The role of B-E in the extinction of an aversive CS+ has been more confusing than that of ACTH/MSH because peripheral administration can itself be aversive (mediated by gut receptors) whereas central injections have the well-documented antianxiety and analgesic properties (Bechara & van der Kooy, 1985). When this distinction is made, it is relatively clear that B-E administered directly into the CNS enhances extinction whereas injections on the other side of the blood-brain barrier retards extinction (Koob, LeMoal, & Bloom, 1984).

This opposed action of ACTH/MSH and B-E in the CNS is also observed in sexual behavior and on a variety of neurobiological measures, including turnover rates of cholinergic and adrenergic neurotransmitters. In addition, there is a well-documented competition between B-E and ACTH for the same targets in the CNS, so called opioid receptors. For instance, an injection shortly before a morphine injection in addicted animals has the same effect as giving the opiate antagonist naloxone (Bertolini & Gessa, 1981; Bertolini, Fratta, Gessa, Montaldo, & Serra, 1984; Bertolini,

Vergoni, Poggioli, & Gessa, 1986; Charney and Redmond, 1983; Jacquet, 1978; Markey & Sze, 1984; Redmond & Huang, 1979; Smock & Fields, 1981). It has also been demonstrated, as indicated in Figure 1, that the influence of a daily pre-session dose of ACTH or VP on extinction is blocked by preventing access to these receptors by a prior injection of naloxone (De Vito & Brush, 1984).

VASOPRESSIN AND OXYTOCIN

The neurohypophysial hormones, vasopressin (VP) and oxytocin, also greatly influence the extinction of conditioning with aversive reinforcers. The effects are also seen with appetitive reinforcers but they are weak and difficult to replicate. Whereas VP reduces extinction, the opposite effect is generally found with oxytocin (Kovacs & Telegdy, 1985; Strupp & Levitsky, 1985; Van Wimersma Greidanus *et al.*, 1983). Several groups have now shown that oxytocin is released from the pituitary as a stress hormone (Verbalis, McCann, McHale, & Stricker, 1986; Williams, Carter, & Lightman, 1985). This release can also be produced by exposure to a Pavlovian CS + for a shock reinforcer (Kelley & Lightman, 1986). Independent of this release from the pituitary, both of these hormones exist in neurons that project from the hypothalamus to a number of limbic sites. There is some evidence for the release of these peptides in the CNS during stress (Laczi, Gaffori, de Kloet, & de Wied, 1983a,b; Laczi, Gaffori, Fekete, de Kloet, & de Wied, 1984).

The large pharmacological doses that were used in many early animal studies with vasopressin were probably themselves an intensely aversive UCS (Ettenburg, van der Kooy, Le Moal, Koob, & Bloom, 1983), however it is doubtful that all of the effects of VP can be attributed to aversive properties of the injections. The strongest evidence for this is that administration of antiserum to VP into the CNS of rats produces the opposite effect to VP on extinction. This is also evidence for VP playing a physiological role in the control of behavior. Similarly, the opposite effects of oxytocin are produced by its antiserum. Moreover, injections of a few nanograms of these hormones directly into limbic structures produces the opposite effects to their antiserum. Finally there are a number of synthetic fragments (such as DG-AVP and DG-LVP) and natural metabolites of these parent hormones that are nearly totally devoid of endocrine effects but that still retain the capacity to influence extinction. Collectively, these studies are powerful evidence for VP and oxytocin being active in the CNS within the physiological range (Burbach, Bohus, *et al.*, 1983; Burbach, Kovacs, & de Wied, 1983b; Van Wimersma Greidanus *et al.*, 1983).

NEUROHORMONES AND THE INCUBATION OF ANXIETY IN ANIMALS

Thus far the discussion has been restricted to consideration of how neurohormones can enhance or retard extinction, that is, how much of a decrement in excitatory strength occurs after the CS + is presented alone. This is not the only possible outcome of unreinforced exposure to an aversive CS +. In some cases the CS + may grow in excitatory strength after exposure. Outcomes of this type are called the incubation

of anxiety (Eysenck, 1979). The understanding of this phenomenon may provide an explanation for the frequent observation that fears in neurotic patients often seem to paradoxically grow worse rather than extinguish in the natural development of these disorders. In the present section, the laboratory evidence will be reviewed that shows that alterations in the levels of certain neurohormones can reliably produce an incubation effect in animals. This evidence is summarized in Table 1.

The influence that brief or extended nonreinforced exposure to a CS+ has on subsequent resistance to extinction can be studied using a three-stage experimental design. In the first stage an excitatory CS+ is established using aversive conditioning; and in the second stage the levels of hormones in the animal are altered and it is then exposed to the CS+ alone. In the third stage, when the hormone levels have returned to normal, the impact of the prior nonreinforced exposure is assessed using an extinction test. If incubation has occurred during Stage 2, then in Stage 3 there should be an increase in the excitatory strength of the CS+. A number of control groups are necessary to make the inference that the critical feature is the contiguity between the elevation in the hormone and exposure to the CS+ in Stage 2. One of the strongest controls is to give a separate group reexposure to the CS+ and also the hormone but at least several hours apart. This balances the exposure to the hormone and CS+ but removes the contiguity between them. When this control

TABLE 1. Influence of Hormones on the Incubation of Anxiety

Group number	Initial training	Subsequent training	Impact on CS
1	CS+ paired with shock	ACTH or epinephrine (EPI) plus 60 sec CS+ exposure	High CS+
2	" "	Saline plus CS+ exposure	Low CS+
3	" "	ACTH or EPI but no CS+ exposure	Low CS+
4	" "	Delayed (5 hour) EPI after CS+ exposure	Low CS+
5	" "	Opioid antagonist then ACTH plus CS+ exposure	Low CS+
6	No training or only CS+ exposure	ACTH or EPI, and/or CS+ exposure	Low CS+
7	Only shocked in a different situation	ACTH or EPI plus exposure to CS+	Low CS+
8	CS+ paired with shock	ACTH plus exposure to CS+ and to novel second-order (S-O) CS	High S-O CS
9	" "	Saline plus exposure to CS+ and novel second-order (S-O) CS	Low S-O CS
10	" "	ACTH plus exposure to novel second-order (S-O) CS but not CS+	Low S-O CS
11	" "	VP plus 5-min exposure to CS+	High CS+
12	" "	VP plus two 5-min exposures to CS+	Very low CS+
13	" "	Saline plus one or two 5-min exposures to CS+	Low CS+

group was used in one incubation study with epinephrine, the paired but not the nonpaired group showed a several-fold increase in fear of the CS+ during the third stage (Kelley, 1985a). Not all studies on the incubation of anxiety have used this control group but a number of other controls have been used (Haroutunian & Riccio, 1977). Incubation effects are also found when ACTH rather than epinephrine is used (Haroutunian & Riccio, 1979). This is not surprising as an acute injection of epinephrine increases ACTH levels in rats (but not humans). However, Haroutunian and Riccio note that they could not obtain the incubation effect when amphetamine was used, thus suggesting that incubation is not attributable to increased arousal *per se*.

In a theoretical discussion of their findings, Riccio and Concannon (1981) suggest that ACTH and epinephrine can produce incubation effects because these hormones are released and become part of the stimuli conditioned during Stage 1 (state-dependent learning); they further suggest that the re-presentation of these stimuli with exogenous injections in Stage 2 reminds the rat of the prior conditioning in Stage 1 and that this is what produces the incubation effect. The difficulty with this explanation is that state-dependent learning does not occur when there are salient spatial cues (a box with a grid floor) that reliably predicts the occurrence of shock; moreover even when these stimuli are not present it takes several-fold more trials than they provided in order for epinephrine to become established as a discriminative CS in instrumental conditioning (Cook, Davidson, Davis, & Kelleher, 1960; Eich, 1980; Jarbe, Svensson, & Laaksonen, 1983). The study by Gray (1975) is often cited as evidence that state-dependent learning can occur with ACTH in a one-trial conditioning; however, given the exceedingly weak shock they used (.13 mA) it is doubtful that any learning occurred at all in this study, and the result is contradicted by other studies. Nevertheless, a more recent experiment by Ahlers and Richardson (1985) showed that blocking the endogenous release of ACTH with dexamethasone given prior to the aversive conditioning stops the influence of an exogenous injection of ACTH during an extinction test some days later; however, as a number of control groups are missing in their experimental design, it is difficult to come to any substantive conclusion from this pilot study.

The effects of peptide fragments related to VP have also been studied using this three-stage experimental design (Krejci, Kupkova, Dlabac, & Kasafirek, 1983). Although the results of this study are difficult to summarize as many different compounds were tested at different dosages, the results in general show that these peptide fragments can also produce incubation. Although the effect was less robust than with the previously cited studies that used either ACTH or epinephrine, this outcome can be attributed to the long CS+ exposure (5 min) given in Stage 2. In fact when the CS exposure period was increased even further (10 min), there was a permanent attenuation of the excitatory strength of the CS+ that was greater than that found in a saline group. This is the opposite of an incubation effect. There are experimental, clinical, and theoretical reasons for believing that incubation effects are dependent on a short-duration CS+ (Eysenck, 1979; Eysenck & Kelley, 1986). In this regard the results of the above study by Krejci *et al.* again support the view that short durations of exposure to the CS+ are critical for demonstrations of incubation; however, the study is also interesting because it suggests VP may not absolutely increase resistance to extinction and that these effects are dependent on the duration of CS+ exposure.

Second-order conditioning is another three-stage experimental design that can be used to investigate instances of hormone-induced changes in excitatory strength of fear cues. In the first stage of this procedure, a CS+ is established by pairing it with a UCS. In the second stage, a new second-order CS is paired with the occurrence of the CS+ but no further shocks are given. The excitatory properties of the second-order CS are then assessed in the third stage. This procedure, by itself, is not very effective for establishing a second-order CS, because the second-order CS is added when the first CS+ is no longer predicting the occurrence of shock; however, if an injection of ACTH or epinephrine is also added to the second stage of this procedure, then the second-order CS acquires a great deal of excitatory strength. Again various control groups are necessary to show that the necessary conditions are the contiguity of the CS+, the second-order CS, and the neurohormone. Although a control group receiving noncontiguous exposure to all three elements has not been supplied, various control groups that received different combinations of just two failed to show the same effect (Concannon, Riccio, & McKelvey, 1980; Concannon, Riccio, Maloney, & McKelvey, 1980).

In the section on extinction, it was noted that blocking the opioid receptors with naloxone stopped the effects of ACTH on extinction; this blockage can also be obtained in the above procedure involving hormone-induced second-order conditioning (Concannon, Riccio, Maloney, & McKelvey, 1980). Again the result demonstrates the critical importance of the opioid receptors. This topic will be discussed again in the section on the physiology of human anxiety but for now it is important to emphasize that these instances of robust second-order conditioning can be taken as another example of hormone-induced changes in excitatory strength of a CS+, and that the incubation of anxiety is a reliable phenomenon. Finally, it is worth underscoring that the results of these second-order conditioning experiments are of potential theoretical and applied importance for explaining the generalization of anxiety of neurotic patients to new stimulus situations. With the appropriate hormonal circumstances, not only can anxiety grow but it can also find new horizons.

NEUROHORMONES, ATTENTION, AND ANXIETY

In the present section, it will be argued that that VP and ACTH/MSH can enhance attention and that it is this property that interferes with extinction and allows the CS+ to acquire a higher level of excitatory strength in an incubation experiment. No attempt will be made here to provide a comprehensive review of neuropeptides and attention; this has been done elsewhere (Beckwith & Sandman, 1978, 1982; Gaillard, 1981; LaHoste, Olson, Kaskin, & Olson, 1980; Pigache & Righter, 1981).

NEUROHORMONES AND ATTENTION IN ANIMALS

One traditional measure of attention is the deceleration in heart rate (bradycardia) during the CS+ in the early stages of conditioning; however, the analysis is confused by the fact that with more intense fear heart rate acceleration is observed.

Two different groups of investigators have shown that ACTH and VP-related peptides delay the extinction of bradycardia to an aversive CS+ (Hagen & Bohus, 1984; Hernandez, 1985). Effects on attention after the administration of ACTH/MSH are also seen in comparisons of performance on intra- and extradimensional shift problems (Beckwith & Sandman, 1978, 1982). A theoretically strong procedure for studying attention in animals is to reduce the capacity of a stimulus to become associated with an UCS by first repeatedly presenting the stimulus alone or random to the occurrence of the UCS and then attempting to convert it into a CS+. Two separate groups of investigators have found that the impact of VP on extinction was only observed in the animals that had been repeatedly exposed to the CS prior to the first conditioning trial; otherwise VP had no effect on extinction (Righter, 1982; Tinus, Beckwith, Wagner, Tinus, & Traynor, 1986). These two studies directly implicate attention, particularly the disinhibition of attention, as one psychological process that VP acts on to produce changes in the performance of conditioned behavior. Considerable evidence from several different groups of investigators suggests that the effects of ACTH, ACTH(4-10) and related peptides are not dependent on prior nonreinforced exposure to the conditioned stimulus (Tinus *et al.*, 1986). This suggests that VP produces its effects on extinction by the disinhibition of attention but that ACTH and related peptides may also have direct excitatory effects on attention.

It has been found that emotionally reactive rats (those that have been bred for higher rates of immobility and defecation in novel situations) are more likely to show incubation effects (Morley, 1977). This suggests that there may be a relationship between environmentally induced changes in emotionality, altered levels of ACTH, and information processing. There is some evidence for this. First, if rats are shocked with very intense electrical shock in one situation and then tested the following day in a different situation (open-field apparatus), the amount of emotionality in the second situation is substantially enhanced compared to nonshock control groups; second, procedures that increase emotionality, such as home-cage crowding also enhance the capacity for stress-induced release of ACTH (Anderson, Crowell, Koehn, & Lupo, 1976; Armario, Castellanos, & Balasch, 1984). These observations, along with the evidence of ACTH influencing attention via the opioid-analgesia receptors, leads to the prediction that if rats were shocked in another environment after receiving nonreinforced exposures to a CS+, then the elevations in ACTH should have a disinhibitory influence on latent inhibition. This enhancement of attention to a weak CS+ has been observed to occur not only with a latent inhibition procedure but also with other accepted procedures for studying attention, such as blocking and overshadowing (Kaspro, Catterson, Schachtman, & Miller, 1984). The previously cited studies thus suggest that a pattern of stressful life events, with a concomitant increase in ACTH production, can seriously change the processing of information by increasing attention to what otherwise might be a weak CS+.

NEUROHORMONES AND ATTENTION IN HUMANS

There is substantial empirical support for the possibility that neurohormones can influence human attention. One group has shown that high circulating levels of ACTH can influence attention (Veith, Sandman, George, & Kendall, 1985). They

utilized patients suffering from congenital adrenal hyperplasia, a disorder in which the patient has high circulating levels of ACTH unless the hypersecretion is controlled by administering cortisone. The patients were tested under the conditions of high and low levels of ACTH, and it was found that the scores on one measure of attention (Sternberg Item Recognition Task) increased or decreased depending on ACTH levels. The same group has independently shown that exogenous injections of ACTH (4-10) will enhance attention in volunteers using this and other measures of attention (Beckwith & Sandman, 1982). Other groups have found similar effects with the structurally related peptide alpha-MSH (LaHoste *et al.*, 1980). Two groups have recently obtained evidence that the pharmacologically potent fragment of ACTH called ORG2766 can have a disinhibitory influence on sustained attention (Born *et al.*, 1984; Fredrickson, d'Elia, & Bengtsson, 1985).

In the process of collecting data on whether VP and related neuropeptides can influence memory, a number of investigators have found evidence that these drugs can influence attention and that changes in attention covary with levels of endogenous VP in some clinical groups (Jennings, Nebes, & Reynolds, 1986; Strupp & Weingartner, 1984; Strupp, Weingartner, Goodwin, & Gold, 1984). In regard to the dishabituating influence of VP on attention, one group has shown that LVP will reduce the progressive attenuation of electrocortical responses that occur during extended recordings of contingent-negative variation (Timsit-Berthier, Mantanus, Devos, & Spiegel, 1982). As this task involves associative conditioning, the results seem to mimic those observed in animals.

In a study with smokers, we have been able to show that individual differences in the functioning of the neurohypophysial system have an influence on sustained attention (Kelley, Lightman, Murphy, & O'Connor, 1986). Because cholinergic agonists are known to be powerful releasers of VP and because difficulties in concentration are a frequent component of the abstinence syndrome in ex-smokers, we reasoned that the variation in occurrence of cognitive complaints may be related to individual differences in the amount of VP released during a nicotine challenge. The abstaining smokers with low postchallenge levels of VP had a greater decline in signal-detection scores and also greater difficulties in self-reported concentration. This provides evidence that physiological differences in the neurohypophysial system are important for attention.

Other evidence shows that the inhibitory effect on attention is produced by changes in levels of oxytocin, opioids, and cortisol (Arsten, Neville, Hillyard, Janowski, & Segal, 1984; Carpenter & Gruen, 1982; Kennett, Devlin, & Ferrier, 1982; Kovacs & Telegdy, 1985; Wolkowitz, Tinklenberg, & Weingartner, 1985). This suggests that there is a parallel between attention and extinction/incubation with respect to the balanced endogenous neurohormonal modulation of behavior, and that any imbalance in this neuromodulation will have consequences for both attention and extinction/incubation. Although the results of these human experiments are often not robust, this is not surprising as the most powerful effects of these neurohormones in animal studies are observed in aversive conditioning experiments; thus, this again implicates the potential role of hormones in anxiety neuroses and human aversive conditioning. Further evidence of these relationships will be considered shortly when the physiology of neuroses is discussed, but first some evidence will be considered that attention is amiss in neurotics.

ATTENTION AND HUMAN ANXIETY

There are a number of reasons for thinking that attention is an important factor in the development of neuroses. One line of evidence comes from the common clinical observation that both anxious and depressed patients complain of difficulties in concentration; and the most tenable explanation of this, at least with the anxious patients, is that fear-related thoughts are intruding into their concentration (Watts & Sharrock, 1985). Recent laboratory studies also support this explanation. For instance, a reduced ability to inhibit orienting to anxiety-related cues has been found to be characteristic of anxious neurotics, whereas normal individuals can direct their attention away from these cues (Macleod, Mathews, & Tata, 1986). This again implicates the absence of inhibitory mechanisms in attention as a operative process in the development of clinical anxiety.

A second line of evidence suggests attention is an important factor in human studies on aversive conditioning. In the research by Öhman and his collaborators, they have been able to demonstrate repeatedly that the types of stimuli to which humans will rapidly form robust conditioned orienting reactions are the same stimuli that are involved in natural phobias. In addition, they have found that the enhanced associability of these potentially phobic stimuli is specific to aversive reinforcers (Öhman *et al.*, 1985). Other research suggests that these instances of selective associability are best accounted for in an attention framework (Kelley, 1986a; Mackintosh, 1983).

There is a basis for thinking that there are two processes involved here, that is, the effects of hormones on attention can be separated from the high associability that allows some stimuli rapidly to acquire excitatory strength. First, the rate of extinction or habituation is in general slower in neurotics and this appears to be independent of the associatively prepared stimulus characteristics. In addition, the monosymptomatic phobics, who are otherwise nonneurotic, show more rapid habituation to all stimuli and also a better response to exposure therapy than patients who are phobic but in general more neurotic (Foa, Stekette, & Ozarow, 1985; Lader, 1980; Sartory, 1983). However, there are also reasons for believing that there may be a synergism between the influence of hormones on attention and the influence of high associability to certain potentially phobic stimuli. Although considerable data from animal studies shows that it is possible to reduce the associability of stimuli by using latent inhibition and learned-irrelevance procedures (Mackintosh, 1983), this is a very difficult phenomenon to demonstrate with a cue that has a high associability with a UCS (Kelley, 1982). The implication for Öhman's conditioning model of neuroses is that the inhibition of attention to potentially phobic stimuli may be very fragile and particularly susceptible to the disinhibitory and excitatory influence of certain neurohormones.

NEUROHORMONES AND HUMAN NEUROSES

Thus far, evidence has been presented that neurohormones can influence the extinction and incubation of aversively conditioned stimuli in animals, that this may be mediated by the influence of hormones on attention, that neurohormones influence

human attention, and that problems in attention are a characteristic of neurotic patients. In the present section, some evidence will be considered that suggests that abnormalities in the neuroendocrine system may contribute to the development of neurotic behavior.

THE INTERPRETATION OF SAMPLES OF STRESS HORMONES

Most of the data that has been collected about hormones and clinical anxiety has been obtained under basal conditions, thus the findings may not be applicable to the functioning of the patient while under stress. Nevertheless, it is well documented that levels of cortisol are often found to be above normal in depressed patients. These elevations correspond with the occurrence of stressful life events. It is however unclear whether these elevations in cortisol itself are contributing to the difficulties in persistence and attention in these patients, but it is also noteworthy that stress down-regulates the corticosteroid receptors in the brains of rats (Dolan, Calloway, Fonagy, De Souza, & Wakling, 1985; Sapolsky, Krey, & McEwen, 1984; Wolkowitz *et al.*, 1985).

Providing neurotic patients with direct exposure to the stimuli they fear and concomitantly obtaining a serial collection of plasma samples can be potentially informative about stress hormones in these disorders. This procedure has been utilized with monosymptomatic phobia patients and it was shown that a modest increase in levels of cortisol occurs after periods of exposure to phobic stimuli; however, some earlier studies failed to find even a small increase (Fredrikson, Sundin, & Frankenhaeuser, 1985; Nesse *et al.*, 1985). Because elevations in cortisol are thought to be important for the inhibition of excessive release of ACTH and CRF, this finding raises the possibility that levels of ACTH and/or CRF may be exceedingly high during exposure to phobic stimuli, thus contributing to the incubation and persistence of phobic fears. Differences in populations aside, this explanation is contradicted by evidence that patients suffering from panic attacks have an above normal secretion of cortisol after an infusion of CRF; however, the same investigators found that basal levels of cortisol and ACTH were higher in these patients. Thus as in depressed patients, this can be taken to possibly mean that CRF levels are higher or more easily elevated in these patients (Roy-Byrne *et al.*, 1986).

There are many disadvantages in using cortisol as a neuroendocrine measure of stress: (a) It is an unreliable index of changes in peripheral ACTH or CRF secretion; (b) it has a low ceiling for showing elevations; and (c) it has a low threshold for release by ACTH. On the positive side, cortisol can be reliably measured from urine and saliva as well as plasma (Fehm, Klien, Holl, & Voigt, 1984; Fehm, Holl, Steiner, Klein, & Voigt, 1984; Raid-Fahmy, Read, Walker, & Griffiths, 1982). There are also problems when levels of plasma ACTH are directly assessed; it is difficult to use these data to make any inferences in regard to increases in ACTH or CRF in the CNS. The evidence is fairly clear that a coupling between central and peripheral levels of stress-related peptide hormones is an untenable assumption, at least in any absolute sense. For instance, social separation in Rhesus monkeys elevates ACTH levels in the cerebrospinal fluid (CSF) but not plasma; and there is very little evidence to suggest that peripheral elevations, at least in the physiological range, influence CSF levels (Kalin, 1986).

It is important also to point out that the absence of CSF elevations of a peptide cannot be taken as evidence that a given hormone or its metabolite is not active in the CNS during stress or that peripheral elevations are not important. A good example here is the recent demonstration that an osmotic challenge with a peripheral injection of hypertonic saline has the same impact on extinction in rats as an injection of VP (Koob, Dantzer, Rodriguez, Bloom, & Le Moal, 1985). This result is particularly striking given that the same procedure fails to elevate CSF levels of VP in rats (Mens, Bouman, Bakker, & van Wimersma Greidanus, 1980). With the earlier cautionary notes in mind, some evidence of possible relationships between neurosis and neuropeptides will now be discussed in more detail.

THE ROLE OF VASOPRESSIN IN NEUROSES

Low levels of VP appear to be associated with depression. After an osmotic challenge, levels of VP in the CSF and plasma are lower in drug-free nonpsychotic unipolar and bipolar depressed patients than in normal subjects. In addition, the administration of antidepressants to these patients reverses this insensitivity to an osmotic challenge, and electroconvulsive shock increases levels of VP (Gold, Ballenger, *et al.*, 1984). Similar findings have been reported by other groups (Gjerris, Hammer, Vendsborg, Christensen, & Rafaelsen, 1985).

Given that depressed patients are noted for their absence of attempts to engage in any adaptive coping behavior and given the substantial animal evidence of a relationship between VP and persistence, the previously cited findings of an inverse relationship between VP and depression have considerable theoretical appeal. Low levels of VP in these patients is also consistent with (a) the observation that the cognitive difficulties in depressed patients are not so much caused by distraction by anxious thoughts as they are by a complete inability to concentrate at all (Watts & Sharrock, 1985), (b) the observation of Kelley *et al.* (1986) that there is an inverse relationship between VP levels after a nicotine challenge and difficulties in concentration and (c) the pharmacological evidence that VP influences attention.

Gold, Ballenger, *et al.* (1984) have shown that depressed patients have low levels of VP after an osmotic challenge, and that manic-depressive patients show swings in levels of VP, which are independent of osmolarity, when they shift between manic and depressive states. In normal groups there is an almost perfect correlation between blood osmolarity and levels of VP. Although one study showed that levels of plasma VP are not altered in acutely anxious nonpsychotic patients, another study has shown that patients suffering from anorexia nervosa had low levels of VP but the dominant characteristic was an exceedingly erratic pattern of VP levels in CSF and plasma after an osmotic challenge (Gold, Kaye, Gary, Robertson, & Ebert, 1983; Raskin, Weitzman, Orenstein, Fisher, & Courtney, 1978). In our research with smokers, we have found an inverse relationship between levels of VP and trait neuroticism on the EPQ, and a positive relationship with the personality trait of psychoticism (Kelley *et al.*, 1986). Collectively these studies suggest that VP levels are often low in neurotics; but more importantly, there is a basis for believing that large swings in levels of VP can occur. The high levels could contribute to the incubation of anxiety by dis-inhibiting attention to potentially phobic stimuli.

THE ROLE OF OPIOID RECEPTORS IN INCUBATION

Earlier, in the sections on incubation and extinction in animals, findings were discussed that show that ACTH and VP competes with the endogenous opioids for occupancy of some of the opioid receptors. In addition, it was observed that these receptors are critical for the hormone-mediated effects on extinction and incubation. It follows that if anxiety neurotics have less central release of opioids, then other peptides might be more likely to occupy these receptors during periods of stress. Figure 2 shows that there is some evidence for this possibility. There is a strong inverse ($-.67$) correlation between levels of total opioid binding in CSF samples and the personality trait of Neuroticism in normal volunteers; moreover an even stronger negative ($-.91$) correlation is found with state anxiety (Post, Pickar, Ballenger, Naber, & Rubinow, 1984). Other investigators have observed a correspondence between low levels of endorphins, low pain thresholds, and high scores on neuroticism (von Kronning, Almay, Johansson, & Terenius, 1978). A third report provides evidence that patients who are neurotic and depressed (i.e., those who show normal suppression of cortisol levels after dexamethasone) have lower beta-endorphin levels in plasma than normal individuals (Cohen, Pickar, Extein, Gold, & Sweeney, 1984). This finding is consistent with the studies discussed earlier showing low cortisol levels in phobics after exposure therapy.

It is likely that alpha-MSH plays a role in the previously noted relationship. First, there is substantial evidence that the release of this peptide in the human CNS occurs during the processing of pro-opiomelanocortin (POMC), which is also the precursor of beta-endorphin, but it is also produced in the CNS apart from POMC. Second, it has similar functional properties to ACTH(4-10) and is an endogenous ligand for the opioid receptors. However, alpha-MSH in the human brain exists in a N-acetylated form, which is behaviorally potent in regard to attention, and also in a desacetylated form, which binds to the opioid receptor involved with analgesia. N-acetylation of beta-endorphin also eliminates its potential to bind to opiate-like receptors, however different enzymes are known to control acetylation for alpha-MSH and beta-endorphin. The importance of this is directly implicated by recent findings suggesting there are differences between depressed and normal individuals in the enzymatic processing of POMC in the CNS, but this possibility has not as yet been separated from the alternative interpretation that there is a difference between normal volunteers and depressed patients in the metabolic clearance of the different peptides. It is of course possible that both possibilities are correct, and either provides a basis by which greater occupancy of the opioid receptors could then occur by ACTH, alpha-MSH or VP. The possibility of individual differences in enzymatic processing or metabolic clearance of these peptides is consistent with the observation that not all subjects show an anxiogenic effect after administration of alpha-MSH (Berettini *et al.*, 1985; Cohen, Pickar, Extein, Gold, & Sweeney, 1984a; Datta & King, 1982; O'Donohue, Handelman, Miller, & Jacobowitz, 1982).

The locus coeruleus is a key structure in the inverse relationship between neuroticism and levels of opioids. It is in this structure that the opioids play a important role in dampening activity in the dorsal adrenergic bundle during stress. Excessive activity in this system can cause profound anxiety. This can be produced by direct

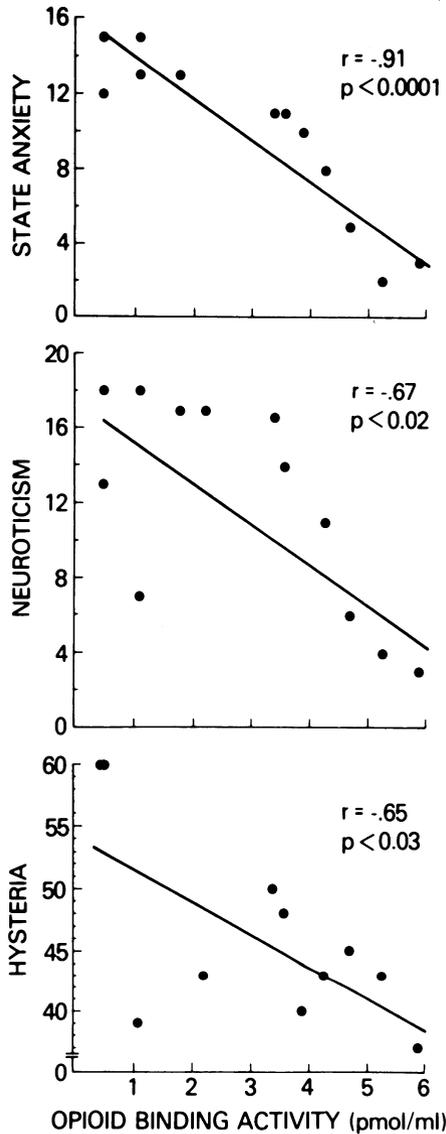


FIGURE 2. An inverse relationship between various measures of anxiety and total opioid binding in CSF samples from volunteers. The figure is reprinted from Post *et al.* (1984).

electrical stimulation of this structure and by abstinence from opiates after tolerance has developed to repeated exogenous injections. The opposite effect can be produced in panic, phobic, depressed, and opiate-withdrawal patients by an infusion of clonidine, but an exacerbation of anxiety occurs after an injection of yohimbine. These effects appear to be mediated, at least in part, by altered activity (Redmond, 1981; Redmond & Huang, 1979). Given this it is not surprising that an abundant literature exists showing relationships between anxiety and various measures of adrenergic functioning (Ballenger, Post, Jimerson, Lake, & Zuckerman, 1984). Finally, a number

of investigators have come to the conclusion that alterations of adrenergic activity in the locus coeruleus have consequences for attention (Aston-Jones, 1985; Robbins, Everitt, & Cole, 1985).

Other neuropeptides are also active in this CNS structure. Although the human data are not yet available, it has been shown with rats that the intraventricular administration of CRF increases the firing of these adrenergic neurons (Valentino, Foote, & Aston-Jones, 1983). A similar involvement is implicated for VP and oxytocin: (a) the greatest extrahypothalamic concentrations of VP and oxytocin exist in the locus coeruleus (Jenkins, Ang, Hawthorn, Rossor, & Iversen, 1984); (b) manipulations that disrupt adrenergic activity interfere with the influence of vasopressin and oxytocin (van Heuven-Nolsen, De Kloet, & Versteeg, 1984; van Wimersma Greidanus *et al.*, 1983); and (c) the development of opioid tolerance can be modulated by these hormones and also another major brain peptide called cholecystokinin (CCK) (Ritzman, Colbern, Zimmermann, & Krivoy, 1984; Vanderhaeghen & Crawley, 1985).

The previously noted influence of VP, CCK, and oxytocin is on the form of opioid tolerance that is subject to associative conditioning, that is, specific to cues that repeatedly predict increased elevations of opioids (O'Brien, Ehrman, & Ternes, 1985; Ritzman *et al.*, 1984). There is evidence that shows that tolerance to endogenous opioids may be involved in the development of clinical depression and neuroticism; but given that most neurotic and depressive reactions are under the control of environmental stimuli, this literature may show some confused findings until this role of stimulus-dependent opioid tolerance is taken into account in the data collection. Nevertheless, in depressive patients who show less than normal suppression of cortisol after taking dexamethasone, there is a greater exacerbation of depressive symptoms after administering high-dose naloxone and also higher levels of beta-endorphin in the CSF (Olson, Olson, & Kaskin, 1985). Also consistent with this possibility are the findings showing that morphine is less effective in suppressing cortisol in these depressed patients than in normal subjects (Zis, Haskett, Albala, Carroll, & Lohr, 1985). There is also reason for believing that the repetitive movements of compulsive rituals may serve the function of enhancing the release of opioids (Cronin, Weipkema, & van Ree, 1986; Henry, 1982). That there is a general similarity between the symptomatology of neurotic anxiety and that of withdrawal from exogenous opiates is also consistent with these possibilities. Finally, as noted earlier phobics have a blunted cortisol response during exposure therapy, and this also occurs in rats made tolerant to exogenous opiates (Buckingham & Cooper, 1984). At the molecular level, it is possible to speculate that after this tolerance to endogenous opioids has occurred, there may be an increased occupancy of the opioid receptor by alpha-MSH or VP during exposure to an aversive CS+.

NEUROHORMONES AND THERAPY FOR NEUROSES

From what has been reviewed thus far, a case can be made that neurohormones are a causal component in the natural development of neurotic behavior. It follows that neurohormones may also be an important consideration in the treatment of these

disorders (de Wied, 1984). Here it is important to point out that the DNA sequences encoding behaviorally active peptides can be genetically cloned (thus inexpensively produced) and that many of these neuropeptides have no classical endocrine effects even when they are conveniently administered orally or intranasally in massive doses (Hayashi, Fischman, Kaskin, & Coy, 1984; van Wimersma Greidanus *et al.*, 1983). This last feature may make these compounds more promising than most of the current anxiolytic drugs on the market. The direct human evidence bearing on these possibilities for such psychopeptidergic therapy will now be considered in the next section. Some of these possibilities may be solely pharmacological interventions that do not take into account behavior therapy at all, being based instead on the drug-conditioning interactions that have been reviewed in the earlier sections, it seems more likely that considerable progress will be made by combining the pharmacological and psychological approaches.

NEUROHORMONES AND THE TREATMENT OF NEUROSES

One possible approach to treating neuroses is a purely pharmacological intervention in which neurohormones or their antagonists might be used directly to alleviate different aspects of neurotic disorders. Given the euphoric properties of opiates, it is not surprising that they have been reported (although never systematically studied) to reduce depression, but this treatment of course has the drawback of the rapid development of tolerance and severe dependence; however, it is now in the realm of possibilities that this drawback may be rectified by concomitant treatment with a compound that reverses tolerance to opiates (Watkins, Kinscheck, & Mayer, 1984). One such compound that has these properties, proglumide, also induces eating in humans and rats, hence it is a possible treatment for anorexia nervosa (Pi-Sunyer, Kissileff, Thornton, & Smith, 1982; Shillaber & Davison, 1984).

A purely pharmacological approach to treating depression has been attempted with a synthetic fragment of VP called DD-AVP. Gold, Ballenger, *et al.* (1984) reported that this fragment produced some but no substantial improvement in depressed patients; however, given (a) that this particular peptide fragment has potent anti-diuretic but no pressor effects and (b) given the evidence of a possible linkage between VP activity at the pressor receptors and the effects on extinction in animals, the null results are subject to an alternative interpretation (Koob, LeBrun, *et al.*, 1985).

BEHAVIOR THERAPY AND NEUROHORMONES

A second approach to the therapeutic use of neuropeptides is to combine their administration with behavior therapy. There are several possible applications. For example, a neurohormone might be used in conjunction with aversive conditioning procedures to enhance the punishment-induced suppression of some undesired behavior or to reduce the extinction of some adaptive defensive behavior. Alternatively, neurohormones might be used with therapies that attempt to extinguish fears by extended exposure to fear-eliciting stimuli. Some examples of these possible applications will now be critically discussed.

Greenberg and Belmaker (1985) recently completed a study with eight patients suffering from agoraphobia and obsessive-compulsive disorders in which DD-AVP was used in conjunction with exposure therapy in a double-blind crossover design. No positive effect of the DD-AVP was observed, but the study is subject to many criticisms. First, the drug was combined with exposure only in one treatment session. Second, as the authors point out, their use of a crossover design is at odds with some evidence that there may be residual effects of VP that were carried over to influence the placebo condition. Nevertheless, even if the sample was cut in half, there were apparently no differences between the four placebo and four DD-AVP subjects when their fears were assessed a week later. Third, as the subjects were instructed to do home-work self-exposure during the intervening week, this may have attenuated the differences between the groups. Fourth, given that the bulk of the evidence suggests that VP reduces extinction, the opposite prediction might have been made concerning the effect of the drug; however, recall that Krejci *et al.* (1983) showed that enhanced extinction occurred in rats provided they received very long exposures to the CS+.

In a second study exploring this combined approach, enhanced passive avoidance was found in three children suffering from Lesch-Nyhan disease when the peptide fragment DD-AVP was administered prior to response-contingent electrical shock (Anderson, David, Bonnet, & Dancis, 1979). This study used a multiple-baseline crossover design in which the children were tested for response to punishment plus saline, to DD-AVP alone, to DD-AVP plus punishment, and on a baseline condition. Suppression of responding was observed only when punishment was combined with DD-AVP; otherwise the punishment contingency had no effect.

A third report involved the intranasal administration of VP or saline to smokers in a randomized two-group design (Ehrensing, Michell, & Baker, 1982). This was done prior to each treatment session in which every inhaled puff of the cigarette smoke was paired with an electrical shock. There was a much greater suppression of smoking in the VP group during the treatment week. Unfortunately, the inference that the drug enhanced the effects of punishment is confounded by the possibility that VP itself might have suppressed the desire for cigarettes (see section on hormones and human attention). The authors report that during the follow-up period the drug group did worse than the control group that received only saline; but this may be a result of the large skewed attrition in the sample size and its composition during the follow-up period, hence the follow-up data are uninterpretable. Nevertheless, the effect during the treatment week supports the results found by Anderson *et al.* (1979).

A fourth study investigated the effect of ACTH (4-10) on aversive conditioning in volunteers (Miller, Fischer, Groves, Rudrauff, & Kaskin, 1977). Although the authors report that they failed to find any influence of the drug on extinction, it is unclear from the sketchy description of the procedure just exactly how the study was performed. In addition, the standard questions raised about any null result in a drug study can be raised about this study. In particular, the critical control is missing for confirmation of a null result, that is, a result showing that the same batch of the drug can otherwise produce a well-documented effect.

Preliminary work with one possible melanocyte stimulating-hormone release-inhibiting factor (MIF-1) suggests that it may be a natural metabolite in the brain and effective in the treatment of depression (Kastin, Olson, Sandman, & Coy, 1981; van der Velde, 1983). This compound has structural (C-terminal amino acid sequence) and functional commonalities with oxytocin, that is, whereas MSH slowed the habituation of the defecation response of rats in repeated trials with the open-field apparatus, MIF-1 had the opposite effect (Datta & King, 1982). Given the intensity of basic pharmacological research on variations of this amino acid sequence, it is likely that several viable compounds will soon become available for testing with humans as an adjunct to behavior therapy (de Wied, 1984; Nicolaidis *et al.*, 1986; Wood *et al.*, 1986).

SUMMARY: TOWARD A CLINICAL NEUROSCIENCE OF ANXIETY

Although it is certainly too early to be able to tell clearly whether behavior therapy will be enhanced by concurrent neurohormonal manipulations, it is clear that the results of a large number of controlled animal studies have provided a firm basis for believing that endogenous hormones are important contributors to resistance to extinction. In addition, we have seen that one of the most perplexing phenomena in clinical neuroses, the incubation of anxiety, appears to be readily explicable as an outcome of an interaction of hormones and conditioning, at least in animals. The next step is to use this knowledge of causation to improve our models of human anxiety by some theoretically focused research on this topic. The research and development for human applications of neuropeptides has thus far largely focused memory disorders; but as I have argued here and elsewhere (Eysenck & Kelley, 1986), the existing evidence suggests that disorders involving attention and anxiety may have an equally important application. In this regard, it will probably not be efficacious to test the therapeutic possibilities of these drugs in clinical populations without first establishing some behavioral assay with volunteers for the effects of these drugs. To accomplish this goal, the human conditioning procedures developed by Öhman *et al.* (1985) may become an important stepping stone to bridge the gap between animal and clinical studies on hormones and anxiety.

REFERENCES

- Ahlers, S. T., & Richardson, R. (1985). Administration of dexamethasone prior to training blocks ACTH-induced recovery of an extinguished avoidance response. *Behavioral Neuroscience*, *99*, 760–764.
- Anderson, D. C., Crowell, C., Koehn, D., & Lupo, J. V. (1976). Different intensities of unsignaled inescapable shock treatments as determinants of non-shock-motivated open field behavior: A resolution of disparate results. *Physiology and Behavior*, *17*, 391–394.
- Anderson, L. T., David, R., Bonnet, K., & Dancis, J. (1979). Passive avoidance learning in Lesch-Nyhan disease: Effect of 1-desamino-8-arginine-vasopressin. *Life Sciences*, *24*, 905–910.

- Armario, A., Castellanos, J. M., & Balasch, J. (1984). Effect of crowding on emotional reactivity in male rats. *Neuroendocrinology*, *39*, 330-333.
- Arnsten, A. F. T., Neville, H. J., Hillyard, S. A., Janowsky, D. S., & Segal, D. S. (1984). Naloxone increases electrophysiological measures of selective information processing in humans. *Journal of Neuroscience*, *4*, 2912-2919.
- Aston-Jones, G. (1885). The locus coeruleus: Behavioral functions of locus coeruleus derived from cellular attributes. *Physiological Psychology*, *13*, 118-126.
- Ballenger, J. C., Post, R. M., Jimerson, D. C., Lake, C. R., & Zuckerman, M. (1984). Neurobiological correlates of depression and anxiety in normal individuals. In R. M. Post & J. C. Ballenger (Eds.), *Neurobiology of mood disorders* (pp. 481-501). Baltimore, MD: Williams & Wilkins.
- Bechara, A., & van der Kooy, D. (1985). Opposite motivational effects of endogenous opioids in brain and periphery. *Nature*, *314*, 533-534.
- Beckwith, B. E., & Sandman, C. A. (1978). Behavioral influences of neuropeptides ACTH and MSH: A methodological review. *Neuroscience and Behavioral Reviews*, *2*, 311-338.
- Beckwith, B. E., & Sandman, C. A. (1982). Central nervous system and peripheral effects of ACTH, MSH, and related peptides. *Peptides*, *3*, 411-420.
- Berettini, W. H., Nurnberger, J. I., Jr., Chan, J. S. D., Chrousos, G. P., Gaspar, L., Gold, P. W., Seidah, N. G., Simmons-Alling, S., Goldin, L. R., Chretien, M., & Gershon, E. S. (1985). Pro-opiomelanocortin-related peptides in cerebrospinal fluid: A study of manic-depressive disorder. *Psychiatry Research*, *16*, 287-302.
- Bertolini, A., & Gessa, G. L. (1981). Behavioral effects of ACTH and MSH peptides. *Journal of Endocrinological Investigation*, *4*, 241-251.
- Bertoloni, A., Fratta, W., Gessa, G. L., Montaldo, S., & Serra, G. (1984). Penile erection during morphine withdrawal: Possible role of ACTH-MSH peptides. In E. E. Muller & A. R. Genazzi (Eds.), *Central and peripheral endorphins: Basic and clinical aspects*. New York: Raven Press.
- Bertolini, A., Vergoni, A. V., Poggioli, R., & Gessa, G. L. (1986). Morphine and B-endorphin antagonist posture and locomotor disorders induced by the injection of ACTH 1-24 in the rat locus coeruleus. *Life Sciences*, *38*, 373-377.
- Bohus, B., de Kloet, E. R., & Veldius, H. D. (1982). Adrenal steroids and behavioral adaptation: Relationship to brain corticoid receptors. In D. Pfaff & D. Ganten (Eds.), *Current topics in neuroendocrinology*, Vol. 2 (pp. 107-151). Berlin: Springer-Verlag.
- Bolles, R. C. (1970). Species-specific defensive reactions and avoidance learning. *Psychological Review*, *77*, 32-48.
- Born, J., Fehm-Wolfsdorf, G., Schiebe, M., Rockstroh, B., Fehm, H. L., & Voigt, K. H. (1984). Dishabituating effects of an ACTH 4-9 analog in a vigilance task. *Pharmacology Biochemistry and Behavior*, *21*, 513-519.
- Breggin, P. R. (1964). The psychophysiology of anxiety. *The Journal of Nervous and Mental Disease*, *139*, 558-568.
- Britton, D. R., Varela, M., Garcia, A., & Rosenthal, M. (1986). Dexamethasone suppresses pituitary-adrenal but not behavioral effects of centrally administered CRF. *Life Sciences*, *38*, 211-216.
- Buckingham, J. C., & Cooper, T. A. (1984). Differences in hypothalamo-pituitary-adrenocortical activity in the rat after acute and prolonged treatment with morphine. *Neuroendocrinology*, *38*, 411-417.
- Buijs, R. M. (1983). Vasopressin and oxytocin—Their role in neurotransmission. *Pharmacology and Therapeutics*, *22*, 127-141.
- Burbach, J. P. H., Bohus, B., Kovacs, G. L., van Nispen, J. W. Greven, H. M., & de Wied, D. (1983). Oxytocin is a precursor of behaviorally active neuropeptides. *European Journal of Pharmacology*, *94*, 125-131.
- Burbach, J. P. H., Kovacs, G. L., & de Wied, D. (1983). A major metabolite of arginine vasopressin in the brain is a highly potent neuropeptide. *Science*, *221*, 1310-1312.
- Cannon, W. B. (1915). *Bodily changes in pain, hunger, fear and rage* (2nd ed.). New York: Appleton.
- Carpenter, W. T., & Gruen, P. H. (1982). Cortisol's effect on human mental functioning. *Journal of Clinical Psychopharmacology*, *2*, 91-101.
- Charney, D. S., & Redmond, D. E., Jr., (1983). Neurobiological mechanisms in human anxiety. *Neuropharmacology*, *22*, 1531-1536.

- Cohen, M. R., Pickar, D., Extein, I., Gold, M. S., & Sweeney, D. R. (1984). Plasma cortisol and Bendorphin immunoreactivity in major and nonmajor depression. *American Journal of Psychiatry*, *141*, 628–632.
- Cohen, M. R., Cohen, R. M., Pickar, D., Sunderland, T., Mueller, E. A., III, & Murphy, D. L. (1984). High-dose Naloxone in depression. *Biological Psychiatry*, *19*, 825–832.
- Concannon, J. T., Riccio, D. C., Maloney, R., & McKelvey, J., (1980). ACTH mediation of learned fear: Blockade by naloxone and naltrexone. *Physiology and Behavior*, *25*, 977–979.
- Concannon, J. T., Riccio, D. C., & McKelvey, J. (1980). Pavlovian conditioning of fear based upon hormonal mediation of prior aversive experience. *Animal Learning and Behavior*, *8*, 75–80.
- Cook, L., Davidson, A., Davis, D. J., & Kelleher, R. T. (1960). Epinephrine, norepinephrine, and acetylcholine as stimuli for avoidance behavior. *Science*, *131*, 990–991.
- Cronin, G. M., Wiepkema, P. R., & van Ree, J. M. (1986). Endorphins implicated in stereotypies of tethered sows. *Experientia*, *42*, 198–199.
- Datta, P. C., & King, M. G. (1982). Alpha-melanocyte-stimulating hormone and behavior. *Neuroscience and Biobehavioral Review*, *6*, 297–310.
- De Vito, W. J., & Brush, F. R. (1984). Effect of ACTH and vasopressin on extinction: Evidence for opiate mediation. *Behavioral Neuroscience*, *98*, 59–71.
- de Wied, D. (1984). CNS effects of oxytocin and related peptides. In P. Paccheri, L. Zichella and P. Falaschi (Eds.), *Endorphins, neuroregulators and behavior in human reproduction* (pp. 4–15). Princeton: NJ: Excerpta Medica.
- de Wied, D., & Jolles, J. (1982). Neuropeptides derived from pro-opiortin: Behavioral, psychological and neuro-chemical effects. *Physiological Reviews*, *62*, 976–1067.
- Dolan, R. J., Calloway, S. P., Fonagy, P., De Souza, F. V. A., & Wakling, A. (1985). Life events, depression and the hypothalamic-pituitary-adrenal axis. *British Journal of Psychiatry*, *147*, 429–433.
- Ehrensing, R. H., Michell, G. F., & Baker, R. P. (1982). Vasopressin's effects on acquisition and extinction of conditioned avoidance response to smoking. *Peptides*, *3*, 527–530.
- Eich, J. E. (1980). The cue-dependent nature of state-dependent retrieval. *Memory and Cognition*, *8*, 157–173.
- Ettenburg, A., van der Kooy, D., Le Moal, M., Koob, J., & Bloom, F. (1983). Can aversive properties of (peripherally-injected) vasopressin account for its putative role in memory? *Behavioural Brain Research*, *7*, 331–350.
- Eysenck, H. J. (1979). The conditioning model of neuroses. *Behavior and Brain Sciences*, *2*, 155–199.
- Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences*. New York: Plenum Press.
- Eysenck, H. J., & Kelley, M. J. (1986). The interaction of neurohormones with Pavlovian A and Pavlovian B conditioning in the causation of neuroses, extinction and incubation of anxiety. In G. Davies (Ed.), *Human conditioning*. London: Wiley.
- Fanslow, M. S. (1985). Odors released by stressed rats produce opioid analgesia in unstressed rats. *Behavioral Neuroscience*, *99*, 589–592.
- Foa, E. B., Stekette, G. S., & Ozarow, B. J. (1985). Behavior therapy with obsessive-compulsives: From theory to treatment. In M. Mavissakalian, S. M. Turner, & L. Michaelson (Eds.), *Obsessive-compulsive disorder*. New York: Plenum Press.
- Fehm, H. L., Klien, E., Holl, R., & Voigt, K. H. (1984). Evidence for extra-pituitary mechanisms mediating the morning peak of plasma cortisol in man. *Journal of Clinical Endocrinology and Metabolism*, *58*, 410–414.
- Fehm, H. L., Holl, R., Steiner, K., Klein, E., & Voigt, K. H. (1984). Evidence for ACTH-unrelated mechanisms in the regulation of cortisol secretion in man. *Klinische Wochenschrift*, *62*, 19–24.
- Frederiksen, S. O., d'Elia, G., & Bengtsson, B.-O. (1985). ACTH 4-9 analogue (ORG 2766) in depressed and elderly patients. *Acta Psychiatrica Scandinavica*, *72*, 349–357.
- Fredrikson, M., Sundin, O., & Frankenhaeuser, M. (1985). Cortisol excretion during defensive reaction in humans. *Psychosomatic Medicine*, *47*, 313–317.
- Gaillard, A. W. K. (1981). ACTH analogs and human performance. In J. L. Martinez, Jr., R. A. Jensen, R. B. Messing, H. Righter, & J. L. McGuagh (Eds.), *Endogenous peptides and learning and memory processes* (pp. 181–196). New York: Academic Press.
- Gjerris, A., Hammer, M., Vendsborg, P., Christensen, N. J., & Rafaelsen, O. J. (1985). Cerebrospinal fluid vasopressin—Changes in depression. *British Journal of Psychiatry*, *147*, 696–701.

- Gold, P. W., Ballenger, J. C., Robertson, G. L., Weingartner, H., Rubinow, D. R., Hoban, M. C., Goodwin, F. K., & Post, R. N. (1984). Vasopressin in affective illness: direct measurement, clinical trials and response to hypertonic saline. In R. M. Post & J. C. Ballenger (Eds.), *Neurobiology of mood disorders* (pp. 323–339). London: Williams & Wilkins.
- Gold, P. W., Chrousos, G., Kellner, C., Post, R., Roy, A., Augerinos, P., Schulte, H., Oldfield, E., & Loriaux, D. L. (1984). Psychiatric and clinical studies of corticotropin releasing factor. *The American Journal of Psychiatry*, *141*, 619–627.
- Gold, P. W., Kaye, W., Robertson, G. L., & Ebert, M. (1984). Abnormal regulation of argine vasopressin plasma and cerebrospinal fluid of patients with anorexia nervosa. *New England Journal of Medicine*, *308*, 1117–1123.
- Gray, P. (1975). Effect of adrenocorticotrophic hormone on conditioned avoidance in rats interpreted as state-dependent learning. *Journal of Comparative and physiological Psychology*, *88*, 281–284.
- Greenberg, D., & Belmaker, R. H. (1985). DDAVP as a possible method to enhance positive benefit of behaviour therapy. *British Journal of Psychiatry*, *147*, 713–715.
- Grossen, N. E., & Kelley, M. J. (1972). Species-specific behavior and acquisition of avoidance behavior in rats. *Journal of Comparative and Physiological Psychology*, *81*, 306–311.
- Hagen, J. J., & Bohus, B. (1984). Vasopressin prolongs bradycardiac response during orientation. *Behavioral and Neural Biology*, *41*, 77–83.
- Haroutunian, V., & Riccio, D. C. (1977). The effect of arousal conditions during reinstatement treatment upon learned fear in young rats. *Developmental Psychobiology*, *10*, 25–32.
- Haroutunian, V., & Riccio, D. C. (1979). Drug-induced “arousal” and the effectiveness of CS exposure in the reinstatement of memory. *Behavioral and Neural Biology*, *26*, 115–120.
- Hayashi, T., Fischman, A. J., Kaskin, A. A., & Coy, D. H. (1984). Some analogs of Try-MIF-1 affect passive avoidance behavior but not motor activity in rats. *Pharmacology Biochemistry and Behavior*, *21*, 809–812.
- Henry, J. L. (1982). Circulating opioids: possible physiological roles in central nervous system. *Neuroscience and Biobehavioral Reviews*, *6*, 229–245.
- Hernandez, L. L. (1985). Vasopressin and ACTH analogs affect orienting but not activity in rabbits. *Peptides*, *2*, 97–101.
- Heuman, K. F. (Ed.). (1985). Proceedings of conference on corticotropin-releasing factor. *Federal Proceedings*, *44*, 145–263.
- Iversen, L. L. (1984). Amino acids and peptides: fast and slow chemicals in the nervous system? *Proceedings of the Royal Society of London*, *221*, 245–260.
- Izquierdo, I., Perry, M. L., Dias, R. D., Souza, D. O., Elisabetsky, E., Carrasco, M. A., Orsingher, O. A., & Netto, C. A. (1981). Endogenous opioids, memory modulation, and state dependency. In J. L. Martinez, Jr., R. A. Jensen, R. B. Messing, H. Rigter, & J. L. McGaugh (Eds.), *Endogenous peptides and learning and memory processes* (pp. 269–304). New York: Academic Press.
- Jacquet, Y. (1978). Opiate effects after adrenocorticotropin or B-endorphin injection in the periaqueductal grey matter of rats. *Science*, *201*, 1032–1034.
- Jarbe, T. U. C., Svensson, R., & Laaksonen, T., (1983). Conditioning of a discriminative drug stimulus: overshadowing and blocking procedures. *Scandinavian Journal of Psychology*, *24*, 325–330.
- Jenkins, J. S., Ang, V. T. Y., Hawthorn, J., Rossor, M. N., & Iversen, L. (1984). Vasopressin, oxytocin and neurophysins in the human brain and spinal cord. *Brain Research*, *291*, 111–117.
- Jennings, J. R., Nebes, R. D., & Reynolds, C. F., III. (1986). Vasopressin peptide (DDAVP) may narrow the focus of attention in normal elderly. *Psychiatry Research*, *17*, 31–39.
- Kalin, N. E. (1986). ACTH in plasma and CSF in the Rhesus monkey. *Biological Psychiatry*, *21*, 124–140.
- Kastin, A. J., Olson, R. D., Sandman, C. A., & Coy, D. H. (1981). Multiple independent actions of neuropeptides on behavior. In J. L. Martinez, Jr., R. A. Jensen, R. B. Messing, H. Rigter, & J. L. McGaugh (Eds.), *Endogenous Peptides and Learning and Memory Processes* (pp. 563–578). New York: Academic Press.
- Kaspro, W. J., Catterson, D., Schachtman, T. R., & Miller, R. R. (1984). Attenuation of latent inhibition by post-acquisition reminder. *Quarterly Journal of Experimental Psychology*, *36B*, 53–63.
- Kelley, M. J. (1982). *Event-reinforcer interactions in conditioning*. Unpublished D. Phil. thesis at The University of Sussex.

- Kelley, M. J. (1985a). Epinephrine-cue contiguity is required for the modulation of CS excitatory strength. *Physiological Psychology*, *13*, 217–218.
- Kelley, M. J. (1985b). Species-typical taxic behavior and event-reinforcer reactions. *Learning and Motivation*, *16*, 301–314.
- Kelley, M. J. (1986a). Selective attention and stimulus-reinforcer interactions in the pigeon. *The Quarterly Journal of Experimental Psychology*, *38B*, 97–110.
- Kelley, M. J., & Lightman, S. L. (1986). *The conditioned release of oxytocin in female rats*. Manuscript submitted for publication.
- Kelley, M. J., Lightman, S. L., Murphy, D., & O'Connor, K. (1986). *Personality, sustained attention and individual differences in nicotine-induced levels of vasopressin*. Manuscript submitted for publication.
- Kennett, D. J., Devlin, M. C., & Ferrier, B. M. (1982). Influence of oxytocin on memory processes: validation by a control study. *Life Sciences*, *31*, 273–275.
- Koob, G. F., LeMoal, M., & Bloom, F. E. (1984). The role of endorphins in neurobiology, behavior, and psychiatric disorders. In C. B. Nemeroff & A. J. Dunn (Eds.), *Peptides, hormones and behavior*. New York: Spectrum.
- Koob, G. F., Dantzer, R., Rodriguez, F., Bloom, F. E., & Le Moal, M. (1985). Osmotic stress mimics effects of vasopressin on learned behavior. *Nature*, *315*, 750–752.
- Koob, G. F., LeBrun, C., Martinez, Jr., J. L., Bluthe, R. M., Dantzer, R., Bloom, F. E., & Le Moal, M. (1985). Use of arginine vasopressin antagonists in eliciting the mechanism of action for the behavioral effects of arginine vasopressin. In R. W. Schrier (Eds.), *Vasopressin* (pp. 195–201). New York: Raven Press.
- Kovacs, G. L., & Telegdy, G. (1985). Role of oxytocin in memory, amnesia and reinforcement. In J. A. Amico & A. G. Robinson (Eds.), *Oxytocin, clinical and laboratory studies* (pp. 359–371). New York: Excerpta Medica.
- Krejci, I., Kupkova, B., Dlabac, A., & Kasafirek, E. (1983). On the effects of neurohypophyseal hormones, analogs and fragments on the extinction of avoidance responding. In E. Endroczi, L. Angelucci, U. Scapagnini, & de Wied, D. (Eds.), *International conference on integrative neurohormonal mechanisms: Neuropeptides and psychosomatic processes*. Budapest: Akademiai Kiado.
- Krieger, D. T. (1983). Brain peptides: What, where, and why? *Science*, *222*, 975–985.
- Laczi, F., Gaffori, O., de Kloet, E. R., & de Wied, D. (1983a). Differential responses in immunoreactive arginine-vasopressin content of microdissected brain regions during passive avoidance behavior. *Brain Research*, *260*, 342–346.
- Laczi, F., Gaffori, O., de Kloet, E. R., & de Wied, D. (1983b). Arginine-vasopressin content of hippocampus and amygdala during passive avoidance behavior in rats. *Brain Research*, *280*, 309–315.
- Laczi, F., Gaffori, O., Fekete, M., de Kloet, E. R., & de Wied, D. (1984). Levels of arginine-vasopressin in cerebrospinal fluid during passive avoidance behavior in rats. *Life Sciences*, *34*, 2385–2391.
- Lader, M. (1980). Psychophysiology studies in anxiety. In G. D. Burrows & B. Davies (Eds.), *Handbook of studies on anxiety* (pp. 59–88). North Holland: Elsevier.
- LaHoste, G. J., Olson, G. A., Kaskin, A. J., & Olson, R. D. (1980). Behavioral effects of melanocyte stimulating hormone. *Neuroscience and Biobehavioral Reviews*, *4*, 9–16.
- Lester, L. S., & Fanslow, M. S. (1985). Exposure to a cat produces opioid analgesia in rats. *Behavioral Neuroscience*, *99*, 756–759.
- Mackintosh, N. J. (1983). *Conditioning and associative learning*. Oxford: Clarendon Press.
- Macleod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, *95*, 15–20.
- Markey, K. A., & Sze, P. Y. (1984). Influence of ACTH in the locus coeruleus of the mouse brain. *Neuroendocrinology*, *38*, 269–275.
- McGaugh, J. L. (1983). Hormonal influence on memory. *Annual Review of Psychology*, *34*, 297–323.
- Mens, W. B. J., Bouman, H. J., Bakker, E. A. D., & van Wimersma Greidanus, T. B. (1980). Differential effects of various stimuli on AVP levels in blood and cerebrospinal fluid. *European Journal of Pharmacology*, *68*, 89–92.
- Meyer, M. E., & Bohus, B. (1983). The modulation of the dorsal immobility response in the adult Wistar rat: The opposite effect of ACTH4-10 and [D-Phe7]ACTH4-10. *Behavior and Neural Biology*, *38*, 194–204.

- Miller, L. H., Fisher, S. C., Groves, G. A., Rudrauff, M. E., & Kaskin, A. (1977). MSH/ACTH influences on the CAR in human subjects: A negative finding. *Pharmacology Biochemistry and Behavior*, *7*, 417–419.
- Morley, S. (1977). The incubation of avoidance behavior: strain differences in susceptibility. *Behavior Research and Therapy*, *15*, 365–367.
- Nesse, R. H., Curtis, G. C., Thyer, B. A., McCann, D. S., Huber-Smith, M. J., & Knoff, R. F. (1985). Endocrine and cardiovascular responses during phobic anxiety. *Psychosomatic Medicine*, *47*, 320–332.
- O'Brien, C. P., Ehrman, R. N., & Ternes, J. W. (1986). Classical conditioning in human opioid tolerance. In S. R. Goldberg & I. P. Stolerman (Eds.), *Behavioral analysis of drug dependence* (pp. 329–356). New York: Academic Press.
- O'Donohue, T. L., Handelman, G. E., Miller, R. L., & Jacobowitz, D. M. (1982). N-acetylation regulates the behavioral activity of alpha-melanotropin in a multineurotransmitter neuron. *Science*, *215*, 1125–1127.
- Nicolaides, E. D., Tinney, F. J., Kaltenbronn, J. S., Repine, D. A., DeJohn, D. A., Lunney, E. A., Roark, W. H., Marriott, J. G., Davis, R. E., & Voigtman, R. E. (1986). Modified di- and tripeptides of the c-terminal portion of oxytocin and vasopressin as possible cognition activation agents. *Journal of Medicinal Chemistry*, *29*, 959–971.
- Öhman, A., Dimberg, U., & Ost, L. (1985). Animal and social phobias: Biological constraints on learned fear responses. In S. Reiss & R. R. Bootzin (Eds.), *Theoretical issues in behavior therapy* (pp. 123–175). New York: Academic Press.
- Olson, G. A., Olson, R. D., & Kaskin, A. J. (1985). Endogenous opiates: 1984. *Peptides*, *6*, 769–791.
- Pi-Sunyer, X., Kissileff, H. R., Thornton, J., & Smith, G. P. (1982). C-terminal octapeptide of cholecystokinin decreases food intake in obese men. *Physiology and Behavior*, *29*, 627–630.
- Pigache, R. M., & Righter, H. (1981). The effects of peptides related to ACTH on mood and vigilance in man. In Tj. B. van Wimersma Griedanus & L. H. Rees (Eds.), *ACTH and LPH in health and disease* (pp. 191–207). Basel: S. Karger.
- Post, R. M., Pickar, D., Ballenger, J. C., Naber, D., & Rubinow, D. R. (1984). Endogenous opiates in cerebrospinal fluid: Relationship to mood and anxiety. In R. M. Post & J. C. Ballenger (Eds.), *Neurobiology of mood disorders* (pp. 356–368). London: Williams & Wilkens.
- Raid-Fahmy, D., Read, G. F., Walker, R. F., & Griffiths, K. (1987). Steroids in saliva for assessing endocrine function. *Endocrine Reviews*, *3*, 367–396.
- Raskin, M. A., Weitzman, R. E., Orenstein, H., Fisher, D. A., & Courtney, N. (1978). Is antidiuretic hormone elevated in psychosis? A pilot study. *Biological Psychiatry*, *13*, 385–391.
- Redmond, D. E., Jr. (1981). Clonidine and the primate locus coeruleus: Evidence suggesting anxiolytic and anti-withdrawal effects. In H. Lal & S. Fielding (Eds.), *Psychopharmacology of clonidine*. New York: Alan Liss.
- Redmond, D. E., Jr., & Huang, Y. H. (1979). New evidence for a locus coeruleus-norepinephrine connection with anxiety. *Life Sciences*, *25*, 2149–2162.
- Riccio, D. C., & Concannon, J. T. (1981). ACTH and the reminder phenomena. In J. L. Martinez, Jr., R. A. Jensen, & R. B. Messing (Eds.), *Endogenous peptides and learning and memory processes*. New York: Academic Press.
- Righter, H. (1982). Vasopressin and memory: The influence of prior experience with the training situation. *Behavioral Neurobiology*, *34*, 337–351.
- Ritzman, R. F., Colbern, D. L., Zimmerman, E. G., & Krivoy, W. (1984). Neurohypophyseal hormones in tolerance and physical dependence. *Pharmacology and Therapy*, *23*, 281–312.
- Robbins, T. W., Everitt, B. J., & Cole, B. J. (1985). Functional hypotheses of the coeruleocortical noradrenergic projection: A review of recent experimentation and theory. *Physiological Psychology*, *13*, 127–150.
- Roy-Byrne, P. P., Uhde, T. W., Post, R. M., Gallucci, W., Chrousos, G. P., & Gold, P. W. (1986). The corticotropin-releasing hormone stimulation test in patients with panic disorders. *American Journal of Psychiatry*, *143*, 896–899.
- Sapolsky, R. M., Krey, L. C., & McEwen, B. S. (1984). Stress down-regulates corticosterone receptors in a site-specific manner in the brain. *Endocrinology*, *114*, 287–292.
- Sartory, G. (1983). The orienting response and psychopathology: Anxiety and phobias. In D. Siddle (Ed.), *Orienting and habituation: Perspectives in human research* (pp. 449–474). New York: Wiley.

- Schacter, S., & Singer, J. (1962). Cognitive social and physiological determinants of emotional state. *Psychological Review*, 69, 379-399.
- Shillabeer, G., & Davison, J. S. (1984). The Cholecystokinin antagonist, proglumide, increases food intake in the rat. *Regulatory Peptides*, 8, 171-176.
- Smock, T., & Fields, H. L. (1981). ACTH 1-24 blocks opiate-induced analgesia in the rat. *Brain Research*, 212, 202-206.
- Strupp, B. J., & Levitsky, D. A. (1985). A mnemonic role for vasopressin: The evidence for and against. *Neuroscience and Biobehavioral Reviews*, 9, 399-411.
- Strupp, B. J., & Weingartner, H. (1984). Vasopressin and cognition. *Drug Development Research*, 4, 501-516.
- Strupp, B., Weingartner, H., Goodwin, F. K., & Gold, P. W. (1984). Neurohypophyseal hormones and cognition. *Pharmacology and Therapeutics*, 23, 179-191.
- Terman, G. W., Shavit, Y., Lewis, J. W., Cannon, J. T., & Liebeskind, J. C. (1984). Intrinsic mechanisms of pain inhibition: Activation by stress. *Science*, 226, 1270-1277.
- Timsit-Berthier, M., Mantanus, H., Devos, J. E., & Spiegel, R. (1982). Action of lysine-vasopressin on human electroencephalographic activity. *Neuropsychobiology*, 8, 248-258.
- Tinus, T. P., Beckwith, B. E., Wagner, N., Tinus, K. A., and Traynor, M. M. (1986). Differential actions of arginine vasopressin and alpha-melanocyte-stimulating hormone on reactivation of memory after hypothermia-induced amnesia. *Physiological Psychology*, 15-22.
- Tomkins, E. H., Sturgis, C. C., & Wearn, J. T. (1919). Studies in epinephrine. II. *Archives of Internal Medicine*, 24, 247-268.
- Vale, W., & Greer, M. (Eds.). (1985). Proceedings of conference on corticotropin-releasing factor. *Federation Proceedings*, 44, 145-263.
- Valentino, R. J., Foote, S. L., & Aston-Jones, G. (1983). CRF activates noradrenergic neurons of the locus coeruleus. *Brain Research*, 363-369.
- Vanderhaeghen, J.-J., & Crawley, J. N. (Eds.). (1985). *Neuronal cholecystokinin: Annals of the New York Academy of Sciences*, Vol. 448. New York: The New York Academy of Sciences.
- van der Velde, C. D. (1983). Rapid clinical effectiveness of MIF-1 in the treatment of major depressive illness. *Peptides*, 4, 297-300.
- van Heuven-Nolsen, D., De Kloet, E. R., & Versteeg, D. H. G. (1984). Oxytocin affects utilization of noradrenaline in distinct limbic-forebrain regions of the rat brain. *Neuropharmacology*, 23, 1373-1377.
- Van Wimersma Greidanus, Tj. B., Bohus, B., Kovacs, G. L., Versteeg, D. H. G., Burbach, J. P. H., & de Wied, D. (1983). Sites of behavioral and neurochemical action of ACTH-like peptides and neurohypophyseal hormones. *Neuroscience and Biobehavioral Reviews*, 7, 453-463.
- Veith, J. L., Sandman, C. A., George, J. M., & Kendall, J. W. (1985). The relationship of endogenous ACTH levels to visual-attentional functioning in patients with congenital adrenal hyperplasia. *Psychoneuroendocrinology*, 10, 33-48.
- Veldhuis, H. D., & de Wied, D. (1984). Differential behavioral actions of corticotropin-releasing factor. *Pharmacology Biochemistry and Behavior*, 21, 707-714.
- Verbalis, J. G., McCann, M. J., McHale, C. M., & Stricker, E. M. (1986). Oxytocin secretion in response to cholecystokinin and food: Differentiation of nausea from satiety. *Science*, 232, 1417-1419.
- von Kronning, L., Almay, B. G. L., Johansson, F., & Terenius, L. (1978). Pain perception and endorphin levels in cerebrospinal fluid. *Pain*, 5, 359.
- Watkins, L. R., Kinscheck, I. B., & Mayer, D. J. (1984). Potentiation of opiate analgesia and apparent reversal of morphine tolerance by proglumide. *Science*, 224, 395-396.
- Watts, F. N., & Sharrock, R. (1985). Description and measurement of concentration problems in depressed patients. *Psychological Medicine*, 15, 317-326.
- Williams, J. L., Drugan, R. C., & Maier, S. P. (1984). Exposure to uncontrollable stress alters withdrawal from morphine. *Behavioral Neuroscience*, 98, 836-846.
- Williams, T. D., Carter, D. A., & Lightman, S. L. (1985). Sexual dimorphism in the posterior pituitary response to stress in the rat. *Endocrinology*, 116, 738-741.
- Wolkowitz, O. M., Tinklenberg, J. R., & Weingartner, H. (1985). A pharmacological perspective of cognitive functions. *Neuropsychobiology*, 14, 133-156.
- Wood, S. P., Tickle, I. J., Treharne, A. M., Pitts, J. E., Mascarenhas, Y., Li, J. Y., Husain, J., Cooper, S.,

- Bludell, T. L., Hruby, V. J., Buku, A., Fischman, A. J., & Wyssbrod, H. R. (1986). Crystal structure analysis of deamino-oxytocin: Conformational flexibility and receptor binding. *Science*, *232*, 633–636.
- Zerbe, R. L. (1985). Genetic factors in normal and abnormal regulation of vasopressin secretion. In R. W. Schrier (Ed.), *Vasopressin* (pp. 213–220). New York: Raven Press.
- Zis, A. P., Haskett, R. F., Albala, A., Carroll, B. J., & Lohr, N. E. (1985). Opioid regulation of hypothalamic-pituitary-adrenal function in depression. *Archives of General Psychiatry*, *42*, 383–386.

CHAPTER 20

Interactions between Drugs and Behavior Therapy

J. A. Gray

Patients given behavior therapy commonly also receive drugs. But neither this fact, nor its potential significance, is likely often to be present in the mind of either the behavior therapist or the prescribing physician. We recently carried out an informal survey at the Institute of Psychiatry to see how far the psychologists treating a group of agoraphobics were aware of the drugs prescribed to their patients. The information they had was rudimentary, though many of their patients were taking, not one, but two or three different compounds. This is hardly surprising, because drugs are typically prescribed by physicians or psychiatrists with little regard for the behavioral treatments that their patients may concurrently receive, and even less discussion with the behavior therapist concerned. Yet, as demonstrated here, there is good reason to suppose that drugs and behavior therapy interact, and not necessarily to the patient's benefit.

This state of affairs is all the more regrettable in that, in the animal laboratory, the discipline of psychopharmacology is perhaps the most thriving offspring yet to emerge from the union of psychology and physiology. Some of the data gathered in psychopharmacological experiments, moreover, suggest certain quite compelling conclusions about likely interactions between drugs and behavior therapy in the clinic. It is the business of this chapter to review some of these conclusions and the data on which they are based.

THE PARTIAL REINFORCEMENT AND PARTIAL PUNISHMENT EFFECTS

Clinical studies suggest that the effective ingredient in most forms of behavior therapy for the phobias is simple exposure to the feared stimulus, therapeutic success being a direct function of the total time that such exposure lasts (Mathews, 1978).

J. A. GRAY • Department of Psychology, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, England.

When we take this into account, two common laboratory phenomena observed most often in rats, the partial reinforcement extinction effect (PREE) and the partial punishment effect (PPE), take on quite good credentials as analogues of behavior therapy.

These two phenomena are most often studied in the simple straight alley, though there is no reason to suppose that they are limited to this apparatus. Indeed, the PREE has been observed in a wide variety of species (including man) and situations; and some of the most important drug effects to be described here have recently been replicated by McNaughton (1984) in the Skinner box. For ease of exposition, therefore, I shall describe a typical experimental paradigm in the alley. The experiment involves two basic behavioral conditions. In the first (continuous reinforcement, CRF) the animal is rewarded (typically, with a pellet of food) on every trial that it traverses the alley. In the second (partial reinforcement, PRF) the animal is rewarded on a randomly chosen proportion (typically, 50%) of trials. Groups of animals are run for equal numbers of trials in the two conditions, and then the running response is extinguished by withdrawal of food reward in both conditions. The PREE consists in the fact that resistance to extinction (measured as number of trials to reach an extinction criterion, or as speed of running to the goalbox during extinction) is reliably greater in the PRF-trained animals.

The PPE is observed in a very similar paradigm, again contrasting two basic behavioral conditions. The first is the same CRF condition as used in the PREE paradigm. In the second, partial punishment (PP) condition, the animals receive food on a randomly chosen proportion of trials, and food preceded by a mild, brief footshock when they enter the goalbox on the remaining trials. The test phase, rather than consisting of extinction as in the PREE paradigm, consists of continuous punishment (food plus shock on every trial) for animals in both groups. The PPE then consists in the fact that PP-trained animals are more resistant to punishment (by the same criteria as in the PREE experiment) than CRF-trained animals.

Now, it is comparatively easy to see how one can treat the PPE as an analogue of clinical exposure therapy. There is little reason to doubt that rats fear shock in much the same way that human phobic patients fear their phobic stimulus (Gray, 1987); and the process by which a rat trained on a PP schedule comes to tolerate shock better than rats not given such previous exposure to shock bears a plausible similarity to the process by which a phobic patient loses his fear in consequence of exposure to phobic stimuli. The case is perhaps less intuitively obvious for the PREE. However, there is considerable evidence that the event called by Amsel (1962) "frustrative nonreward" (i.e., the nonoccurrence of an expected reward) has the same aversive properties as a painful stimulus such as an electric shock. Behaviorally, the two types of event have similar consequences; hormonally, they both cause the release of stress hormones (corticosteroids); and neurologically they are largely mediated by overlapping brain structures (Gray, 1975, 1982, 1987). Furthermore, it is often (Brown & Wagner, 1964; Chen & Amsel, 1977), though not always (J. Feldon, personal communication), possible to demonstrate cross-tolerance between punishment and nonreward: that is, initial exposure to footshock gives rise, not only to increased resistance to punishment with further shock, but also to increased resistance to extinction; and initial exposure to nonreward gives rise to increased resistance to

punishment as well as extinction. This phenomenon suggests, not only that nonreward bears important similarities to physical pain, but also that exposure to either can under some circumstances engage a rather general process of increased resistance to stressors of many kinds. Other experiments have demonstrated a similar cross-tolerance between electric shock and the stress of forced swimming in cold water (Weiss, Glazer, & Pohorecky, 1976). The process that gives rise to this general tolerance for stress has been called, by Miller (1976), "toughening up." The task of the behavior therapist, and the lives of his patients, might be made considerably easier if we knew how to harness this process.

THE EFFECTS OF ANTIANXIETY DRUGS ON THE PREE AND PPE

Either the PPE or the PREE, then, may plausibly be used as an analogue of the effects of behavior therapy. The series of experiments I go on to describe has made use of both of these phenomena. These experiments have been concerned with the influence on the PREE and PPE of drugs that are used clinically to control anxiety. Nowadays, these drugs are drawn largely from the family of benzodiazepines or from a range of newer drugs that bind to the same receptor as the benzodiazepines. Before the discovery of the benzodiazepines in the early 1960s, the drugs used for this purpose tended to be barbiturates or the carbamate, meprobamate. Earlier still (and for some of us even today) the drug of choice was alcohol. Animal experiments show that, although these classes of drugs are chemically diverse, they act in much the same way on behavior (Gray, 1977), the chief difference being that the margin between a selectively anxiolytic dose and a dose at which sedative or other unwanted effects prevail is larger for benzodiazepines, giving these compounds their therapeutic superiority over the earlier drugs. The experiments described here have made use of the benzodiazepine, chlordiazepoxide (Librium), or the barbiturate, sodium amylobarbitone (Sodium Amytal).

The main clinical question to which one would like to know the answer is this: If a patient is receiving behavior therapy for the treatment of his phobias, is it useful, harmful, or indifferent for him also to be treated with anxiolytic drugs? On the assumption that the PREE and PPE are indeed valid analogues of the effects of behavior therapy, the animal experiments described in the following suggest that the answer to this question may depend on at least the following factors: (a) the type of behavior therapy; (b) the intervals between exposures to feared stimuli; (c) the timing of drug administration with respect to exposure; and (d) the presence or absence of the drug at the time the effects of exposure are measured. It will be easier to understand the roles played by these factors if we first outline the theoretical analysis of the PREE that has emerged from three decades of research in the animal laboratory (Mackintosh, 1974).

Two theories of the PREE (due to Amsel, 1962, and Capaldi, 1967, respectively) have received substantial empirical support. Fortunately, the two theories are not incompatible with one another and turn out each to deal best with different ranges of data. It is possible, therefore, to combine the two theories into one reasonably comprehensive and cogent account of the phenomenon.

Consider the sequence of events experienced by a rat trained on a PRF schedule. On the first few rewarded trials it learns that running to the goalbox secures a food reward. Once this expectation of reward is established the absence of food in the goalbox on a nonrewarded trial is able to act as the aversive event of frustrative nonreward, eliciting the emotional state that Amsel (1962) has called primary or unconditioned frustration.

There is much evidence that the state of unconditioned frustration is functionally equivalent to, and perhaps mediated by the same brain systems as, the emotional state elicited by unconditioned painful stimuli (Gray, 1967, 1987; Wagner, 1966). But there is also much evidence that the state elicited by such unconditioned aversive events, and the physiological system mediating it, are different from the state and physiological system brought into play by secondary or conditioned stimuli that have been associated (by Pavlovian conditioning) with primary aversive events (Gray, 1982, 1987). In the context of the PREE we may term these states unconditioned and conditioned (or anticipatory; Amsel, 1962) frustration respectively. In a more general context I have suggested that conditioned frustration is but one manifestation (the conditioned fear elicited by stimuli associated with pain being another; Mowrer, 1960) of the emotion of anxiety; and that the system (the behavioral inhibition system) that mediates this state consists of a set of interconnected brain structures centered on the septohippocampal system (Gray, 1982). The brain structures that mediate the behavioral effects of unconditioned aversive stimuli are quite separate; they include the amygdala, portions of the hypothalamus, and the central gray of the midbrain (Adams, 1979; Graeff, 1987; Gray, 1987; Panksepp, 1982). The emotion (if any) to which activity in this system (the fight/flight system; Gray, 1971, 1987) corresponds is not well specified (rage and terror have about equally good claims).

With these more general distinctions in mind, let us return to our rat being trained on a PRF schedule in the alley. Such a rat will frequently experience a rewarded trial occurring after a nonrewarded trial at a time determined by the intertrial interval (ITI). According to Capaldi's (1967) analysis of the PREE, it is these nonreward-to-reward (N-R) transitions that give rise to the increased resistance to extinction shown by a PRF- relative to a CRF-trained animal. Briefly, Capaldi proposes that the PRF-trained rat continues to experience a trace of the nonrewarded trial at the start of the rewarded trial and so comes to learn that performance of the instrumental response (running down the alley) in the presence of this trace leads to reward. The trace was at first treated by Capaldi as an immediate aftereffect of nonreward; this might be identical to Amsel's unconditioned frustration reaction, although Capaldi himself left it unspecified. Later, he preferred to speak of the memory of the nonrewarded trial; such a memory could include all the various features, physical as well as emotional, that differentiate an empty goalbox from one that contains food reward. Whether aftereffects or memories are involved, the associations so formed ensure that the sequence of nonrewarded trials that constitutes extinction produces a number of cues that, for the PRF- but not the CRF-trained animal, elicit continued running, so giving rise to the PREE.

Amsel (1962) postulates a further process that can give rise to the PREE. Once unconditioned frustration comes to be elicited by nonreward, it can be conditioned (like any other unconditioned response) to stimuli that regularly precede nonreward.

Such stimuli are provided by the startbox and stem of the alley and by the act of running down it. These, then, come to elicit conditioned frustration. It follows that, on rewarded trials, the PRF-trained animal will be rewarded for running in the presence of internal stimuli that characterize the state of conditioned frustration (stimuli that, according to my generalization of Amsel's theory, are identical to those that characterize the state of anxiety). In this way, just as an animal trained in this way comes to run in the presence of Capaldi's aftereffects or memories of nonreward, so will it learn to run in the presence of the internal cues of conditioned frustration. For the CRF-trained animal, unconditioned and conditioned frustration will occur for the first time during extinction and will rapidly (owing to their intrinsically aversive character) disrupt running behavior. The PRF-trained animal, however, has learnt to persist when it experiences conditioned frustration, and so will run longer during extinction—the PREE.

These two accounts of the PREE are clearly not incompatible with each other. And, indeed, the evidence is that both Capaldi's and Amsel's process contribute to the PREE (Mackintosh, 1974). However, the contributions that each makes differ depending on the exact experimental parameters used. The critical parameters are themselves predictable from the two theories.

Capaldi's process depends on the animal's carrying over a trace of one (nonrewarded) trial to the next (rewarded) one. Whether this trace is an aftereffect or a memory, it would be expected to lose strength as the ITI grows longer (though the time course of this loss would be expected to be much briefer for aftereffects than for memories). In contrast, Amsel's process depends on Pavlovian conditioning, which is relatively insensitive to the intervals between trials; and, once the conditioned frustration response is established, its elicitation is completely independent of the ITI. This analysis leads to the prediction that the PREE should be relatively more dependent on Capaldi's process at short ITIs, a prediction that is well supported by experimental findings (Mackintosh, 1974).

A second important variable is the number of acquisition trials. Amsel's process depends on the buildup of Pavlovian conditioned frustration, and this should become stronger with repeated experiences of nonreward. Capaldi's process, in contrast, depends on the formation of associations between particular nonrewarded trials and following rewarded trials. (For readers with an interest in cognitive psychology, it is perhaps worth likening this distinction to that between semantic and episodic memory, the former corresponding to Amsel's process, the latter to Capaldi's.) This analysis leads to the prediction that the PREE should be relatively more dependent on Amsel's process after many acquisition trials. This prediction, too, is supported experimentally (Mackintosh, 1974).

There is one further difference between Amsel's and Capaldi's processes that is worth bearing in mind when we come to consider the effects of anxiolytic drugs upon the PREE. Capaldi's process gives rise to associations between the act of rewarded running (or other instrumental behaviour), on the one hand, and relatively specific stimuli emanating from a nonrewarded trial, on the other. An Amselian rat, in contrast, forms an association between a very general emotional state (conditioned frustration, or the still more general state of anxiety; Gray, 1982) and the act of rewarded running. It follows that the persistence that depends on Capaldi's process

should be more severely weakened by changes between the conditions of acquisition and extinction (e.g., in the nature of the apparatus or the instrumental response) than the persistence that depends on Amsel's process. Although no direct experimental comparison bearing upon this inference has, to my knowledge, been made, the results of a number of experiments are consistent with it (e.g., Amsel & Rashotte, 1969).

Armed with this theoretical apparatus, let us turn to the effects on the PREE of the anxiolytic drugs. I have described Amsel's conditioned frustration as a species of the more general state of anxiety. If so, one might expect that antianxiety drugs would weaken conditioned frustration. And, indeed, there is a wealth of evidence that antianxiety drugs antagonize the behavioral effects of stimuli associated with either nonreward or punishment (but not the effects of unconditioned nonreward or punishment), a state of affairs that one may summarize by saying that these agents impair the functioning of the behavioral inhibition system (Gray, 1977, 1982).

Applied to the PREE, this generalization gives rise to the prediction that anxiolytic drugs should impair this effect especially at long ITIs (when the contribution of Capaldi's process should be relatively small). This prediction is well supported by the data. If the ITI is set at 24 hours, both sodium amylobarbitone (Feldon, Guillaumon, Gray, De Wit, & McNaughton, 1979) and chlordiazepoxide (Feldon & Gray, 1981) are able to block the PREE completely, the partially reinforced animals trained under the drug extinguishing like CRF-trained animals. At an ITI of about 5 minutes, in contrast, blockade of the PREE occurs largely in a state-dependent manner (Gray, 1969; Ison & Pennes, 1969; McNaughton, 1984), a point to which we return later. The same pattern of results emerges from experiments on the PPE: at a 24-hour ITI, chlordiazepoxide given during training abolishes the PREE; at a 5-minute ITI, it is without effect (Davis, Brookes, Gray, & Rawlins, 1981).

A further prediction is that, at short ITIs, blockade of the PREE by anxiolytic drugs should be greater, the larger the number of acquisition trials (because the contribution of Amsel's process grows with length of acquisition). This prediction, too, is supported by the data, though less clearly: there is no non-state-dependent effect of sodium amylobarbitone on the PREE at 7 (Ziff & Capaldi, 1971) or 48 (Ison & Pennes, 1969) acquisition trials, but some impairment at 72 trials (Gray, 1969).

But these are not the only effects that we would expect the anxiolytics to exert on the PREE. Recall that Capaldi's process requires the animal to form a specific association between the events that constitute a nonrewarded trial (or the episodic memory of this trial) and rewarded instrumental behavior. It follows that any change between the conditions accompanying the nonrewarded trials of acquisition and extinction, respectively, should weaken the contribution of Capaldi's process to the PREE. Amsel's process, in contrast, depending as it does on an association between an internal state and rewarded instrumental behavior, should be more resistant to the effects of such changes. Now one source of such changes is the drug state itself. Suppose we train an animal on a PRF schedule while it is drugged but extinguish the instrumental response in the absence of the drug. To the extent that Capaldi's process contributes to the PREE, the change from drug to no drug between acquisition and extinction should correspondingly weaken the PREE. This kind of state-dependent

impairment can be distinguished from a direct blockade of the PREE by an examination of the condition in which animals are trained on a PRF schedule without the drug but extinction is conducted under the drug: state dependence should be a symmetrical phenomenon, so that any loss of the PREE in groups trained under drug and extinguished without the drug should be mirrored by an equivalent loss in groups trained without the drug but extinguished with it.

More precise prediction is hampered by the fact that the drug-nonreward associations on which state dependence (of the kind relevant to this discussion) depends are likely themselves to vary as a function of experimental parameters, including those (ITI and number of acquisition trials) that affect the relative predominance of Amsel's and Capaldi's processes in the determination of the PREE. Furthermore, state dependence is known to vary according to the particular drug used. Thus we can do little more than summarize the brute experimental facts pertaining to the effects of state dependence on the PREE. At a short ITI, effects of this kind have been reported with 48 (Ison & Pennes, 1969) and 72 (Gray, 1969; McNaughton, 1984) but not 6 (Ziff & Capaldi, 1971) acquisition trials; all these experiments used sodium amylobarbitone. At a 24-hour ITI, no state dependent effects on the PREE were seen with chlordiazepoxide (Feldon & Gray, 1981). Using sodium amylobarbitone, however, Feldon *et al.* (1979) observed different effects depending on the segment of the alley in which measurements were made. In the goal segment, the drug in training abolished the PREE directly, there being no state dependent loss of the PREE. In the start section of the alley, only state dependency was observed: the PREE was absent in both conditions in which drug treatment was switched between acquisition and extinction but present in the conditions in which amylobarbitone was administered in both acquisition and extinction or in neither. The intermediate, run section of the alley showed both effects, the PREE being abolished when amylobarbitone was given during acquisition irrespective of drug treatment during extinction, and also in the groups given no drug during training but amylobarbitone during extinction. This pattern of results may indicate that retrieval of episodic memory (Capaldi's process) affects performance preferentially at the initiation of the instrumental response, at least when long ITIs are used.

CLINICAL IMPLICATIONS

It is time to see how these experimental findings bear on the clinical issues to which the argument of this chapter is ultimately directed. The experiments on the PREE and the PPE described above suggest the following conclusions: (a) Under some conditions anxiolytic drugs are able directly and completely to block the process by which exposure to an anxiogenic event creates behavioral tolerance for that event; these conditions include long intervals between exposure to such events and large numbers of such exposures. (b) As a separate effect, anxiolytic drugs are also able to limit the expression of behavioral tolerance for an exposed anxiogenic event to the state in which the subject is in receipt of the drug; this state dependence occurs if exposure is conducted under the drug, and is a function of parameters of the exposure regime that are at present poorly understood. (c) State dependence may also take

the form that tolerance for anxiogenic events that is acquired and expressed in the absence of drug treatment is lost when the subject is given an anxiolytic drug.

It will be clear that any of these effects, if they occur clinically, would be inimical to the patient's recovery. Furthermore, effects (a) and (b) can co-exist, as in Feldon *et al.*'s (1979) observations in the run section of the alley (see earlier). Note that the clinical hazard represented by state dependency (effect b) is no less than that represented by direct blockade of the development of tolerance for stress (effect a). Indeed, the hazard represented by effect b may even, from one point of view, be considered the greater. A patient in whom this effect is of significant magnitude would find, when taken off anxiolytic medication, not only an acute rise in the level of anxiety, but also a loss of the tolerance for anxiety that had been acquired under the drug. In this way, the combination of anxiolytic medication and behavior therapy might contribute to dependence on continued medication. A similar outcome might occur in the absence of formal behavior therapy. It is well known that the spontaneous recovery rate among patients suffering from neurotic complaints is about 70% (Eysenck, 1952). This recovery rate may reflect the operation under normal life circumstances of the same processes that are put deliberately to work by the behavior therapist. If so, it follows that, even in the absence of formal behavior therapy, anxiolytic medication may prevent behavioral recovery or limit recovery to the drugged state, along the lines described earlier.

If we search the existing clinical literature for evidence relating to these deductions, there is little useful information. Such formal studies of interactions between anxiolytics (all benzodiazepines) and behavior therapy as have been conducted (Sartory, 1983) have involved only limited durations of combined drug and behavioral treatment (days or weeks) and the period of follow-up has similarly been limited. But in normal clinical practice it is common for both behavioral and pharmacological treatments to continue for many weeks or months. Even given these limitations in the existing data base, there is some indication of clinical deterioration in patients given drugs during behavioral treatment as compared to those given behavioral treatment only, particularly at the longer times of follow-up (Hafner & Marks, 1976; Sartory, 1983). This, of course, is as we would predict from the animal experiments. More anecdotal evidence comes from Sartory's (personal communication) studies of long-term users of benzodiazepines suffering from agoraphobia. She has found that such patients apparently benefit less from behavioral treatments than do nonusers of anxiolytic drugs. Inevitably, there are many interpretations that could be put on such observations. It may be, for example, that individuals who are predisposed to become dependent on benzodiazepines have personality traits that also render them unsuitable for behavior therapy, or that they are simply more seriously ill than nonusers of these drugs. But a further possibility, consistent with the arguments advanced earlier, is that the use of benzodiazepines hampers behavior therapy, and/or that prior use has prevented the spontaneous development of any useful degree of tolerance for anxiogenic stimuli, so increasing the dropout rate when exposure to such stimuli is employed in therapy.

Given the paucity of clinical evidence, we have recently begun a direct study of the interactions between a benzodiazepine—diazepam (Valium)—and *in vivo* exposure in the treatment of agoraphobia. Patients entering the study will be of two kinds,

long-term (at least 3 months) benzodiazepine users and nonusers, randomly allocated to drug or placebo medication but all given *in vivo* exposure over a period of 7 weeks. Follow-up will last a year. We anticipate that this study will allow us to evaluate the degree to which concomitant benzodiazepine treatment alters the course and effectiveness of exposure therapy for agoraphobia.

A MORE POSITIVE NOTE

So far the stance I have adopted in this discussion has been, therapeutically speaking, defensive: I have considered the possibility that the type of combination between drug and behavioral treatments that is commonly (but inadvertently) found in current clinical practice is harmful or, at any rate, less than optimal. However, other lines of argument may lead in more positive directions.

To begin with, the evidence from experiments with animals suggests that behavioral tolerance for stress may be developed in more than one way, and that not all of the processes involved are affected by the anxiolytic drugs. We have already seen that the PREE depends on (at least) two processes, Amsel's and Capaldi's, and that anxiolytics apparently affect directly only Amsel's. A third process, different from both of these, by which behavioral tolerance for aversive stimulation may be increased is that of Pavlovian counterconditioning (Dickinson & Pearce, 1977).

Pavlovian counterconditioning depends on an association in which an aversive stimulus (e.g., electric shock) serves as the conditioned stimulus for an appetitive unconditioned stimulus (e.g., food). As a consequence of the establishment of this association the aversive stimulus comes to elicit conditioned responses appropriate to the appetitive stimulus (e.g., salivation) and, at the same time, to lose some of its aversive characteristics. It is the latter effect that is of interest here. A clinical procedure whose effects may reflect the same underlying process is that of systematic desensitization (Wolpe, 1958). In this procedure, presentation of an anxiogenic event (usually, though not necessarily, in imagination) is immediately followed by an event of more positive hedonic tone, sometimes taking the form of a tangible physical or social reward, sometimes that of instructed relaxation. To the extent that this pairing between anxiogenic and hedonically positive events is responsible for the patient's loss of the initial anxiety response (and this extent remains a matter of controversy; Mathews, 1978), we are probably dealing with Pavlovian counterconditioning.

If this analogy between Pavlovian counterconditioning and systematic desensitization is sound, it becomes important to know whether anxiolytic drugs affect counterconditioning. We have conducted one experiment addressed to this question, and obtained a clear answer to it. In this experiment (McNaughton & Gray, 1983) rats were first trained to press a bar for food reward delivered on an intermittent schedule. When they had established a steady rate of barpressing, an association was set up between a tone and a shock, both delivered while the rat was barpressing. In consequence the tone came to suppress the rate of barpressing (a so-called conditioned emotional response). At this point in the experiment a further complication was introduced. During occasional intrusion periods, lasting a minute at a time, the bar was withdrawn from the experimental chamber and one shock and one food reward

were delivered to the animal independently of its behavior. For one (control) group of rats the shock and the food delivery were randomly related in time; for the other, counterconditioning group, the food always followed immediately after the shock. The effect of this Pavlovian pairing between shock and food was that the tone (which was still, as before, followed by the shock in both groups) came to suppress barpressing less in the counterconditioning group than in the control group, demonstrating that Pavlovian counterconditioning (i.e., a reduction in the aversive properties of the shock) had indeed taken place. Furthermore, the magnitude of this counterconditioning effect was in no way reduced by the administration of an anxiolytic dose of the benzodiazepine, chlordiazepoxide.

If this result can be generalized to other instances of counterconditioning, and if systematic desensitization is such an instance, we may infer that systematic desensitization should not be affected by the concomitant administration of benzodiazepines. This inference may offer a way of helping the long-term user of benzodiazepines in need of behavioral treatment. It is difficult to persuade such patients to tolerate the exposure to anxiogenic stimuli, which forms a necessary part of all behavioral treatments for anxiety, especially if they are first taken off drugs (Sartory, personal communication). Yet, as we have seen, there is reason to suppose that administration of anxiolytic drugs may interfere with what is currently the most commonly used behavioral treatment for phobias and obsessions, namely, *in vivo* exposure. One avenue of treatment that may be worth exploring, therefore, is that of conducting systematic desensitization while the patient remains under anxiolytic medication. This, too, is an issue that we are investigating in our current clinical study: long-term benzodiazepine users will be randomly allocated to either *in vivo* exposure or systematic desensitization under either placebo or diazepam drug treatment.

A still more hopeful possibility emerges from a recent report by Shemer, Tykocinski, and Feldon (1984) that behavioral tolerance to punishment and nonreward is proactively increased if rats are given a course of chlordiazepoxide injections before any behavioral training commences. In this experiment the animals were first injected with chlordiazepoxide daily for 12 days, then trained to run in the alley for food reward, and finally tested either in extinction or by the addition of punishment to reward. Controls were treated identically, except that initially they received a course of placebo injections. The animals that had been given chlordiazepoxide (several weeks prior to the test phase of the experiment) were more resistant to both punishment and extinction. Neither the psychological nor the physiological mechanism of action of this unusual drug effect is yet known. It is nonetheless still possible to raise the question whether similar effects might take place in man. If so, the optimum time to administer behavioral treatments, such as *in vivo* exposure, might be just after the patient has terminated a course of anxiolytic medication. We shall have the opportunity to make a preliminary investigation of this hypothesis in the study of long-term benzodiazepine users described earlier, because for some of these patients drug administration will be terminated just before behavioral treatment commences. However, a more adequate test of the hypothesis will require administration of a benzodiazepine (or placebo) to nonusers of the drug for a period of time that terminates before the start of behavior therapy.

Notice that the findings reported by Shemer *et al.* (1984) suggest that behavioral tolerance for stress may be induced by purely pharmacological means. Stone (1979), indeed, suggested that it is precisely in this way that antidepressant drugs work. The essentials of his argument are as follows.

It is known that exposure to stressors of many different kinds leads to the release in the brain of monoamine neurotransmitters, especially noradrenaline. One consequence of this activation of noradrenergic neurons is that, if stress is severe and/or prolonged, there is an acute fall in the amount of transmitter available in the neuron for further release; a second consequence is that receptors postsynaptic to the emitting neuron are exposed to high levels of stimulation. Neurochemical adaptation to both these effects is known to take place (Fillenz, 1977). To counteract the fall in the level of the neuron's stores of noradrenaline, there is an increase in the amount and/or activity of the enzyme, tyrosine hydroxylase, which is rate limiting in the synthesis of the transmitter. At the same time, to compensate for the increased stimulation of postsynaptic noradrenergic receptors, these undergo so-called down-regulation, that is, a reduction in number and/or affinity for the transmitter.

There is some evidence that the former effect (increased activity of tyrosine hydroxylase) underlies at least one instance of behavioral tolerance for stress. Weiss *et al.* (1976) report that one session of inescapable shock disrupted the rat's subsequent capacity to escape or avoid shock in a shuttlebox, but that this disruptive effect was lost after 14 consecutive daily sessions of inescapable shock (toughening up). Furthermore, after one but not 14 sessions of inescapable shock, there was a fall in the level of noradrenaline in the brain; and after 14 but not one session of shock, there was a rise in the activity of tyrosine hydroxylase in the brain, suggesting that this constitutes the neurochemical basis of toughening up. This is a possibility to which we return later. Stone (1979), however, concentrates on the second, receptor-based adaptation to stress-induced release of noradrenaline. He pointed out that antidepressant drugs generally have two neurochemical effects in common: acutely, they increase the concentration of monoamines, especially noradrenaline, in the synapse; and, chronically, they cause down regulation in receptors postsynaptic to the monoamine transmitter, especially in β -noradrenergic receptors. Stone proposed, therefore, that the therapeutic effects of the antidepressants are due to their mimicking the normal synaptic consequences of exposure to stress, so giving rise to the neurochemical process that underlies the development of tolerance for stress. This hypothesis predicts that a course of antidepressant medication, given before a rat is tested with shock or in extinction (as in the Shemer *et al.* experiment described earlier), should lead to increased resistance to punishment or extinction. It similarly predicts that such a course of medication, given before the start of behavior therapy, should facilitate its effects. I know of no data relevant to either of these predictions.

The possibility of increasing behavioral tolerance for stress by purely physical means is also raised by some recent experiments from my own laboratory. These experiments derive from the observations, described earlier, that anxiolytic drugs impair the PREE and PPE.

For some years we have been trying to work out the route of action by which the anxiolytic drugs produce these effects (Gray, 1982). Our experiments have focused

on the septohippocampal system and the noradrenergic neurons that originate in the locus coeruleus (in the brainstem) and innervate this system (along with many other regions of the forebrain). The hippocampal formation displays a characteristic electrographic pattern, known as the theta rhythm: high voltage, regular, almost sinusoidal slow waves (in the range 6-12 Hz) that can easily be recorded from free-moving small mammals (rats, guinea pigs, rabbits, cats, dogs, etc.) via electrodes chronically implanted in the brain. The theta rhythm is controlled by pacemaker cells (almost certainly cholinergic) located in the medial septal area and that project their axons into the hippocampal formation. Via electrodes chronically implanted in the septal area it is possible to drive, or block, the hippocampal theta rhythm by electrical stimulation: a train of low-intensity (c. 60 μ A), short-duration (c. 0.5 msec) pulses with frequencies within the normal theta range drives the theta rhythm at the imposed frequency; a train of high-frequency pulses (above about 70 Hz) blocks the spontaneous theta rhythm for the duration of the stimulation. It is possible to plot a function that relates the frequency of the septal driving current to the threshold current capable of eliciting the theta rhythm. When this is done in the free-moving male rat, there is a characteristic minimum in the resulting curve, located precisely at 7.7 Hz (inter-pulse interval, 130 msec). This minimum in the theta-driving curve is due to the action of noradrenergic neurons reaching the septohippocampal system, because destruction of these neurons abolishes the minimum while leaving the basic generation of the theta rhythm intact. The same effect is produced by systemic administration of any of a range of anxiolytic drugs, including benzodiazepines, barbiturates, alcohol, and meprobamate, suggesting the hypothesis that these compounds exert their influence on behavior by blocking the noradrenergic input to the septohippocampal system and the facilitation of 7.7 Hz theta that this input entails. This hypothesis is supported by a number of observations showing that stress increases activity in noradrenergic neurons and that this increased activity is reversed by all classes of anti-anxiety drugs.

Suppose, then, that the anti-anxiety drugs block the PREE and PPE by way of the neurophysiological route just described. One might then expect to be able to produce effects opposite to those caused by these compounds if one were artificially to drive the hippocampal theta rhythm at that frequency, namely 7.7 Hz, which is selectively antagonized by the drugs. In an initial test of this hypothesis I showed that it is indeed possible to increase resistance to extinction (so producing a pseudo-PREE) by using septal stimulation to drive theta at 7.7 Hz on a random 50% of occasions when a rat, having traversed an alley, is in the goalbox drinking a water reward (Gray, 1972). These results were consistent with the hypothesis that 7.7-Hz theta is an internal signal of conditioned frustration, as defined by Amsel (1962), and that my experimental paradigm permitted associative counterconditioning between this signal and water reward. However, subsequent research has shown that the elicitation of 7.7-Hz theta causes behavioral tolerance for stress by a simpler mechanism. For, in the relevant experiments (Holt & Gray, 1983a, b, 1985) we have applied septal theta-driving stimulation before any behavioral training takes place, so eliminating the possibility of associative counterconditioning.

The basic design used in these experiments contrasts two groups of rats, one given a course of 10 days' septal theta-driving stimulation (about 90 seconds a day),

the other treated identically except that no current is passed via the implanted electrodes. Subsequent to this stimulation phase of the experiment, both groups are trained to press a bar for food reward or a fixed-ratio schedule. Finally, we attempt to eliminate barpressing by subjecting it to extinction, punishment with footshock, or conditioned suppression by a tone paired with footshock. The stimulated group shows increased resistance to suppression of barpressing in all three of these procedures relative to the controls. Two features of our results suggest, furthermore, that they are indeed due to the driving of theta that the septal stimulation produces (as distinct from the many other neural pathways that this stimulation undoubtedly also activates). First, high-frequency, theta-blocking septal stimulation causes the reverse effect, proactively decreasing resistance to extinction. Second (J. Williams, personal communication), the proactive increase in resistance to extinction caused by low-frequency septal stimulation is critically dependent on stimulation frequency, occurring only at frequencies including and close to 7.7 Hz (the exact details of this frequency dependence are still subject to investigation).

These results suggest, therefore, that elicitation of the hippocampal theta rhythm at specific frequencies lying close to 7.7 Hz engages a nonassociative process, similar to that described by Miller (1976) as toughening up (see the description of the Weiss *et al.* experiments, earlier), and giving rise proactively to increased behavioral tolerance for stress. A further similarity between our results and those reported by Weiss *et al.* (1976) came to light when we measured the activity of tyrosine hydroxylase in the hippocampus of our stimulated rats: relative to the unstimulated controls, those animals that showed increased resistance to extinction also showed elevated tyrosine hydroxylase activity (Graham-Jones, Holt, Gray, & Fillenz, 1985). This effect may be produced at sites within the hippocampus where the septohippocampal cholinergic projection innervates noradrenergic presynaptic terminals (Fung & Fillenz, 1983). If so, there are striking parallels between the neurochemical elements that participate in adaptation to stress in the sympathetic nervous system and adrenal cortex, on the one hand, and the septohippocampal system, on the other. For the chain of events by which stress causes an acute fall in the levels of noradrenaline, followed by an increase in the activity of tyrosine hydroxylase and the eventual restoration of normal levels of noradrenaline, was first described in these peripheral organs; and there the increase in tyrosine hydroxylase activity is known to depend on the cholinergic synaptic input to noradrenergic sympathetic nerves and adrenal chromaffin cells (Axelrod & Reisine, 1984).

There is a long way to go before the details of the neurochemical changes by which the brain adapts to stress (so causing behavioral adaptation to stress) are worked out. But research of this kind holds out the promise that we shall one day possess a chemical means of inducing behavioral tolerance for stress to supplement existing behavioral methods. For this promise to be realized, it is vital, I believe, that pharmacologists and psychologists should pursue their approaches to the study and treatment of neurosis in tandem rather than, as so often at present, in mutual ignorance or even rivalry. In clinical parlance today the term *psychopharmacology* usually means no more than the selection of a drug from the psychiatrist's shelf. It is time it came to represent a true fusion between two disciplines, each of which needs the other.

REFERENCES

- Adams, D. B. (1979). Brain mechanisms for offence, defense, and submission. *Behavioral and Brain Sciences*, 2, 201–241.
- Amsel, A. (1962). Frustrative nonreward in partial reinforcement and discrimination learning: Some recent history and a theoretical extension. *Psychological Review*, 69, 306–28.
- Amsel, A., & Rashotte, M. E. (1969). Transfer of experimenter-imposed slow response. *Journal of Comparative and Physiological Psychology* 69, 185–189.
- Axelrod, J., & Reisine, T. D. (1984). Stress hormones and their interaction and regulation. *Science*, 224, 452–459.
- Brown, R. T., & Wagner, A. R. (1964). Resistance to punishment and extinction following training with shock or non-reinforcement. *Journal of Experimental Psychology*, 68, 503–507.
- Capaldi, E. J. (1967). A sequential hypothesis of instrumental learning. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation* (Vol. 1, pp. 67–156). New York: Academic Press.
- Chen, J. S., & Amsel, A. (1977). Prolonged, unsignalled, inescapable shocks increase persistence in subsequent appetitive instrumental learning. *Animal Learning and Behavior*, 4, 377–385.
- Davis, N. M., Brookes, S., Gray, J. A., & Rawlins, J. N. P. (1981). Chlordiazepoxide and resistance to punishment. *Quarterly Journal of Experimental Psychology*, 33B, 227–239.
- Dickinson, A., & Pearce, J. M. (1977). Inhibitory interactions between appetitive and aversive stimuli. *Psychological Bulletin*, 84, 690–711.
- Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology*, 16, 319–324.
- Feldon, J., & Gray, J. A. (1981). The partial reinforcement extinction effect after treatment with chlordiazepoxide. *Psychopharmacology*, 73, 269–275.
- Feldon, J., Guillaumon, A., Gray, J. A., De Wit, H., & McNaughton, N. (1979). Sodium amylobarbitone and responses to nonreward. *Quarterly Journal of Experimental Psychology*, 31, 19–50.
- Fillenz, M. (1977). The factors which provide short-term and long-term control of transmitter release. *Progress in Neurobiology*, 8, 251–278.
- Fung, S. C., & Fillenz, M. (1983). The role of pre-synaptic GABA and benzodiazepine receptors in the control of noradrenaline release in rat hippocampus. *Neuroscience Letters*, 42, 61–66.
- Graeff, F. G. (1987). The anti-aversive action of drugs. In T. Thompson, P. B. Dews, & J. Barrett (Eds.), *Advances in behavioral pharmacology* (Vol. 6). Hillsdale, NJ: Erlbaum.
- Graham-Jones, S., Holt, L., Gray, J. A., & Fillenz, M. (1985). Low-frequency septal stimulation increases tyrosine hydroxylase activity in the hippocampus. *Pharmacology, Biochemistry & Behaviour*, 23, 489–493.
- Gray, J. A. (1967). Disappointment and drugs in the rat. *Advancement of Science*, 23, 595–605.
- Gray, J. A. (1969). Sodium amobarbital and effects of frustrative non-reward. *Journal of Comparative and Physiological Psychology*, 69, 55–64.
- Gray, J. A. (1971). *The psychology of fear and stress*. London: Weidenfeld & Nicolson; New York: McGraw-Hill.
- Gray, J. A. (1972). Effects of septal driving of the hippocampal theta rhythm on resistance to extinction. *Physiology and Behavior*, 8, 481–490.
- Gray, J. A. (1975). *Elements of a two-process theory of learning*. London: Academic Press.
- Gray, J. A. (1977). Drug effects on fear and frustration: Possible limbic site of action of minor tranquilizers. In L. L. Iversen, S. D. Iversen, & S. H. Snyder (Eds.), *Handbook of psychopharmacology* (Vol. 8, pp. 433–529). New York: Plenum Press.
- Gray, J. A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. Oxford: Oxford University Press.
- Gray, J. A. (1987). *The psychology of fear and stress* (2nd ed.). Cambridge: Cambridge University Press.
- Hafner, J., & Marks, I. (1976). Exposure *in vivo* of agoraphobics: Contributions of diazepam, group exposure, and anxiety evocation. *Psychological Medicine*, 6, 71–88.
- Holt, L., & Gray, J. A. (1983a). Septal driving of the hippocampal theta rhythm produces a long-term, proactive and non-associative increase in resistance to extinction. *Quarterly Journal of Experimental Psychology*, 35B, 97–118.
- Holt, L., & Gray, J. A. (1983b). Proactive behavioural effects of theta-blocking septal stimulation in the rat. *Behavioral and Neural Biology*, 39, 7–21.

- Holt, L., & Gray, J. A. (1985). Proactive behavioural effects of theta-driving septal stimulation on conditioned suppression and punishment in the rat. *Behavioral Neuroscience*, *99*, 60–74.
- Ison, J. R., & Pennes, E. S. (1969). Interaction of amobarbital sodium and reinforcement schedule in determining resistance to extinction of an instrumental running response. *Journal of Comparative and Physiological Psychology*, *68*, 215–219.
- Mackintosh, N. J. (1974). *The psychology of animal learning*. New York: Academic Press.
- McNaughton, N. (1984). Effects of anxiolytic drugs on the partial reinforcement extinction effect in runway and skinner box. *Quarterly Journal of Experimental Psychology*, *36B*, 319–330.
- McNaughton, N., & Gray, J. A. (1983). Pavlovian counterconditioning is unchanged by chlordiazepoxide or by septal lesions. *Quarterly Journal of Experimental Psychology*, *35B*, 221–233.
- Mathews, A. (1978). Fear-reduction research and clinical phobias. *Psychological Bulletin*, *85*, 390–404.
- Mowrer, O. H. (1960). *Learning theory and behavior*. New York: Wiley.
- Miller, N. E. (1976). Learning, stress and psychosomatic symptoms. *Acta Neurobiologiae Experimentalis*, *36*, 141–56.
- Panksepp, J. (1982). Towards a general psychobiological theory of emotions. *Behavior and Brain Science*, *5*, 407–467.
- Sartory, G. (1983). Benzodiazepines and behavioral treatment of phobic anxiety. *Behavioural Psychotherapy*, *11*, 204–217.
- Shemer, A., Tykocinski, O., & Feldon, J. (1984). Long term effects of chronic chlordiazepoxide administration. *Psychopharmacology*, *83*, 277–280.
- Stone, E. A. (1979). Subsensitivity to norepinephrine as a link between adaptation to stress and antidepressant therapy: an hypothesis. *Research Communications in Psychology, Psychiatry and Behavior*, *4*, 241–55.
- Wagner, A. R. (1966). Frustration and punishment. In R. M. Haber (Ed.), *Current research on motivation* (pp. 229–239). New York: Holt, Rinehart, Winston.
- Weiss, J. M., Glazer, H. I., & Pohorecky, L. A. (1976). Coping behavior and neurochemical changes: An alternative explanation for the original “learned helplessness” experiments. In A. Serban & A. Kling (Eds.), *Animal models in human psychobiology* (pp. 141–173). New York: Plenum Press.
- Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.
- Ziff, D. R., & Capaldi, E. J. (1971). Amytal and the small trial partial reinforcement effect: stimulus properties of early trial nonrewards. *Journal of Experimental Psychology*, *87*, 263–269.

PART VI

CONCLUSION

CHAPTER 21

Concluding Comments on Theoretical Foundations and Requirements in Behavior Therapy

Irene Martin

INTRODUCTION

Since its inception behavior therapy has continuously examined its own theoretical foundations and shifting orientations, and has questioned their relationship to clinical practice. In the same spirit, contributors to this book were asked to present their views on contemporary theory and its relevance to behavior therapy. They represent a spectrum of researchers ranging from those engaged in day-to-day therapy to those whose primary concern is with fundamental issues in theory. Hence the chapters reflect a number of different attitudes to theory and a number of different perspectives on what are seen as the requirements of an adequate theory. Chapter 1 affirms the starting point of Behavior Therapy as rooted in Pavlovian conditioning theory. It might be expected that succeeding chapters would show how these origins developed along the lines laid down. This is not the case. Rather, they illustrate the diversity of contemporary developments. Today's conditioning research refers to many issues, to cognitive representations, to information processing, to learning about the causal structure of the world, and to emotions and the influence of neurohormones on acquisition and extinction.

Such diversity does not imply the presence of discrepant and conflicting theories, but testifies to the multifaceted components of the conditioning process. The answer to the simple question of what is learned in conditioning paradigms turns out to involve an elaborate group of processes that wait to be integrated within a more comprehensive theoretical framework of learning.

IRENE MARTIN • Department of Psychology, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, England.

Chapters in the cognitive section similarly illustrate a heterogeneous range of models and theories. Discussed here are network theories, mental models, attributional styles and schemata, and a number of themes relevant to therapy yet in a preliminary stage of analysis: the cognition–emotion interface, conscious and unconscious information processing, the structural architecture of internal representations. The experimental cognitive laboratory research has its own distinctive set of paradigms that examine different aspects of behavior and from a different perspective, typically that of inferring structure from what is observed. It too is in a fragmented state and offers no comprehensive theory.

Clinical practice and interests have themselves diversified. The original concern with overt behavioral difficulties exemplified in the various phobias has moved to depression, panic, anxiety states, and to less well-defined problems. Within this broader spectrum of interests patterns of maladaptive behavior are seen to be diffuse, less easily described or classified. Patients seen in the clinic present a relatively arbitrary constellation of symptoms possessing little in the way of homogeneity, and an important issue is how to subdivide such heterogeneity into therapeutically relevant groupings. Behavior therapists have expressed discontent with DSM-III classifications as being arbitrary and only poorly related to treatment. They have always championed a functional analysis of behavior in preference to psychiatric categories, and have aimed to link such analyses with specific intervention procedures. One contemporary view is that this kind of analysis remains useful but needs to be supplemented by analysis of cognitive misperceptions and misconstructions.

As a result of all these factors—the heterogeneity of fundamental theoretical research, its lack of obvious relevance to practice and the change in orientation within behavior therapy—there has been a call to abandon the old theoretical formulations and look for new ones. Such a call has to be resisted. The fact that no comprehensive conditioning or cognitive theory is available may make it difficult to keep up with basic research and the changing nature of its questions and foci. But many fundamental principles remain and there is no good reason to ignore the effective foundations that have already been laid down, or to ignore the success of other disciplines faced with the same problem of bridging the gaps between theory and practice.

Eysenck reminds us that this kind of situation is characteristic of scientific development in its early stages, where paradigms are typically seen as being inadequate, crude, limited in scope, and having little predictive power (Eysenck, 1985). The fact that a large number of questions remains concerning the application of learning theory to behavior therapy is no argument for abandoning it. Contemporary conditioning research is in a lively state, and although it is seen by many clinicians as not making specific enough connections with the day-to-day requirements and language of clinical practice, there is evidence of considerable effort being made along these lines in several of the chapters.

Cognitive psychologists are aware that their models deal almost exclusively with the higher cognitive functions of semantic analysis, meaning, and reasoning, and have little to say about emotions and maladaptive behavior. As a result, clinicians have tended to generate theories within the clinical context that seem to make much more immediate sense and to offer useful practical guides. Beck's schemata, Lang's information-processing theory of emotion, and Leventhal's schema construct, among

others, refer to cognitive structures that involve expressive-motor reactions, memories, images, and ideas, that is, the whole range of processes that are relevant to maladaptive behavior and that lie outside the models and paradigms of academic cognitive psychology. These newer models carry the implication that therapeutic interventions can be achieved in a number of ways—via cognitive restructuring, extinction of conditioned responses, direct behavioral skills training—manipulations that will lead to behavior change.

The emergence of these hybrid models illustrates how material can be drawn from different approaches and traditions—cognitive and conditioning, psychophysiology and semantics, thinking and feeling—and integrated within a comprehensive system with clearly specified aims. These are (a) to incorporate both conditioning and cognitive mechanisms of maladaptive behavior; (b) to encompass larger-scale patterns of behavior rather than the usual specific and isolated components of laboratory studies; and (c) to relate (a) and (b) to therapeutic goals.

The implicit assumption, which will not be questioned here, is that it is behavior change that forms the goal of behavior therapy: whatever the nature of the model employed and whatever the therapeutic technique applied, the end product is a change in observed behavior. To achieve some existential change in viewing life and its purpose is not considered to be within the realm of behavior therapy. From this perspective we can consider how the theories discussed in the chapters relate to overt behavior and principles of behavior change.

THE CONDITIONING VIEW OF BEHAVIOR

The simplest hypothesis is that behavior is under stimulus control: present the stimulus and some unit of behavior will inevitably follow. In such an account no mediating knowledge structure is necessary, although from the start conditioning theory has used mediating concepts in the form of a variety of intervening variables and hypothetical constructs. As the conditioning chapters amply demonstrate, accusations that conditioning theory is reflexological, mechanistic, stimulus bound, and that it fails to take individual strategies into account are now outdated.

They illustrate how much conditioning research and theory have changed in recent years, and the heterogeneous nature of the contemporary framework. Today's conditioning research encompasses silent learning, registration of stimulus features, representational structures, and hence its relationship to observed behavior in particular has changed a great deal. Researchers are no longer exclusively concerned with the description of functional stimulus–response laws.

Critical theoretical issues in the contemporary approach to classical conditioning require the use of probabilistic and/or inhibitory operations or contingencies. Probabilistic associations of CS/UCS occurrence or absence leads to learning, but this learning is not necessarily manifested in an overt change in an external response. Although the immediate effect of learning can only be inferred from a change in overt behavior, many cognitively oriented learning psychologists describe what is learned in terms of predictive aspects of the environment rather than in terms of behavior change. This approach is illustrated in Dickinson's chapter, where he refers to the function of conditioning as enabling animals to detect causal relations between

events in their environment, with the implication that the laws of conditioning will tend to mirror the laws of causality.

Other conditioning researchers have explored different avenues. Perhaps one of the most immediately relevant is Susan Mineka's work on fear conditioning in monkeys (Chapter 4). Behavior as described here is unambiguously fearful and includes a number of well-defined behavioral components of fear. Her research shows that acquisition of fear can occur in a very few trials, it examines the effectiveness of various extinction procedures, and has obvious and useful analogues to behavior therapy.

Research into fear and anxiety has frequently turned to physiological evidence, and this is illustrated in Kelley's chapter, which reviews recent animal work on endogenous hormones and their influence on behavior in aversive conditioning paradigms. He also examines evidence that neurohormones influence human attention, thereby offering a link with clinical research that suggests that anxious people are unable to inhibit orienting to anxiety-related cues.

Research explicitly concerned to link animal-conditioning research with clinical issues is described in Gray's chapter, which poses a specific, clinically relevant question: If a patient is receiving behavior therapy for the treatment of a phobia, is it useful, harmful, or indifferent for him or her also to be treated with anxiolytic drugs. Such research is directed toward analyzing the nature of the interaction between pharmacological agents and kinds of behavioral treatment in order to determine their optimal combination in therapy, as well as the further goal of how they might both be effectively used to bring about behavioral tolerance to stress.

Another question that has received attention is whether conditioned behavior is driven by a cognitive plan (e.g., Breger & McGaugh, 1965). One view of human conditioning states that behavior is not entirely or even mainly stimulus driven but that it follows from rules that individuals can learn, can verbalize, and can then use to guide behavior (Brewer, 1974). This extreme view has not been substantiated by the evidence, and most conditioning researchers accept that although the individual's knowledge and hypotheses may influence responding to some degree, there are limits to the cognitive control of conditioned behavior. The extent to which such rule learning determines performance in classical conditioning is examined in Chapter 6 where the evidence shows that under certain conditions, individuals' use of a rule that is imparted either by instructions or learned through experience does significantly affect responding by inhibiting it on unreinforced trials and facilitating it on reinforced trials. Nevertheless, the data of such experiments show that some individuals who can verbalize the contingencies do not or cannot use such information to govern their responding. Being able to verbalize a rule does not guarantee matching behavior.

Lowe *et al.* discuss the emergence of rule-governed behavior in the context of operant conditioning. They cite evidence that the behavior of adult humans on basic schedules of reinforcement differs considerably from animal performance under similar conditions. These distinctive features of adult operant behavior are absent in preverbal infants, who perform on some schedules in a manner indistinguishable from that of animals. Again the data show marked individual differences. These authors suggest that token economy programs might be more effective if more attention were given to the therapist's explicit verbal description of the contingency in planning reinforcement schedules. Therapeutic applications of conditioning that do

not consider the controlling role of verbal behavior will inevitably be inadequate; verbal control forms an integral part of token economy programs and must be taken into account.

The factor of verbalizable knowledge about stimulus events is seen as an important contributor to responding in both classical and operant conditioning experiments, though the extent of such control and the mechanisms by which verbal knowledge interacts with unconscious knowledge or stimulus-driven behavior remains an empirical issue.

The evidence as it stands suggests that stimulus control of conditioned responding does occur under some conditions, and that the terms *automatic* and *unmediated* can reasonably be employed to convey the directness of these stimulus-response connections. Under other conditions, the effects of knowledge about contiguity, contingency, patterning of events, and awareness about the response itself, seem to provide additional control over conditioned responding.

Eifert discusses the role of verbalizations, with the reminder that verbal stimuli themselves can acquire affective/emotional properties by means of classical conditioning. In this way humans develop a verbal-emotional repertoire consisting of a large number of words that come to be emotional stimuli. If complex language is one of the key characteristics that distinguish humans from animals, what are the behavioral-directive functions relevant to behavior therapy?

Direct therapeutic use of language occurs in positive self-reinforcement and self-instructional statements, and in semantic desensitization procedures applied to phobias. Eifert illustrates the explicit use of language conditioning in self-instructional training programs in which clients are taught to emit self-statements that are incompatible with and opposite in emotional content to the negative self-statements they have previously employed. Like other contributors to this section, he stresses that at the present time the precise relationships between different types of self-statement and overt behavior are not clearcut. Performance-based treatment methods are significantly more effective in producing behavior change than methods that rely solely on verbal or imagined procedures, and maintenance of behavior change is enhanced by combining performance-based procedures with verbal techniques, such as self-instruction training.

So far as conditioning and behavior are concerned, it seems that the part of conditioning dealing with emotional reactions and overlearned habits continues to assume some degree of stimulus control of behavior. In other areas, such an assumption does not have central importance. Fundamental animal conditioning research uses change in overt behavior to infer the structure of what has been learned, and interest in describing the specifics of behavior has declined to make way for interest in the kinds of inferences that animals make about their world.

So far as human conditioning is concerned, interest has shifted from the stimulus-response functions that govern conditioned performance to the ways in which cognitive factors influence it, illustrated for example in subjects' verbalizations about their knowledge. Subjects who verbalize rules of reinforcement behave differently from those who do not. However, explanations of these effects are lacking, and it is evident that verbalization is itself very complex and interpreted in different ways. In one sense it can be an indicator of the subject's knowledge. In the sense that Eifert

uses it, it refers to a verbal-emotional repertoire consisting of words that act like emotional stimuli.

Overall, the conclusion has to be reached that conditioning can occur outside awareness. We can suppose that preverbal children, like other animals, will rapidly acquire a range of behavioral responses, with the implication that some forms of behavior—evaluations, emotions, psychophysiological and interoceptive responses—are laid down early in life. Such primitive forms of conditioning persist throughout life in situations that involve emotion and demand rapid action. With experience in the world and long-term exposure to event sequences, the individual learns to assess correlations and contingencies that, with the development of language, can be formally expressed. The extent to which stimulus-driven behavior interacts with a more cognitively determined learning to result in the behavior patterns of everyday life is one of the major questions underlying clinical theory and practice.

A point to emphasize is that none of the different points of view discussed by the conditioning contributors excludes the others. All are legitimate ways of explaining what is learned in conditioning. Thus *conditioning* today is an umbrella term that describes factors relating to survival action, to longer-term regulation of regularities in event sequences, and estimates of probability of event occurrence; it includes changes in central representations, hormonal changes, and all such factors contribute to overt behavior. We do not have to choose between viewing conditioning in one or other of these terms.

As Dickinson points out, contemporary theory provides a more liberal model of conditioning than the traditional one, but it is also a more complex and uncertain one. Whereas behavior therapists of the previous generation could treat conditioning as a simple and well-understood phenomenon, they will now find in the literature a plethora of effects describing what is learned in the conditioning situation. The important question—How does what is learned affect overt behavior?—remains largely open. Most contemporary conditioning theories, like the stimulus–stimulus theories of old, are still faced with the task of explaining how an association between two stimuli or between their centers or representations generates a change in behavior. The simple answer that the CS activates the representation of the reinforcer and hence the same set of responses, is inadequate. It fails to account for response development, the growth of adaptive or maladaptive responding, skilled performance, or the complex behavioral diversity of emotional responses and their maintenance.

The best that can be said at present is that behavior that is persistent and inappropriate to our current goals suggests the operation of a primitive conditioning mechanism. Such responses—generally referred to as habits—imply that the control of action has become independent of our knowledge about its consequences and as a result, autonomous of the current value of the goal or reinforcer, a form of behavioral autonomy (Dickinson, Chapter 3).

COGNITIVE MODELS

Clinical accounts of the cognitive learning approach assume that internal events are better predictors of human behavior than external variables. Response output is not considered to be a direct function of the physical input, but a function of intervening

cognitive events. However, no single theoretical development characterizes this area, and the fragmentary and paradigm-bound nature of cognitive theory has been generally recognized and widely discussed.

Contributors tackle the difficult problem of applying cognitive models to clinical practice in a number of ways. Available models can be roughly classified as structural or functional. The former emphasizes the organization and structure of mental processes. Information-processing models fall within the functional variety, an approach illustrated in Eysenck and Mathews' discussion of ways in which anxious individuals process threatening or threat-related stimuli.

Structural models are discussed by Power. He examines the nature of the structural organization in such models as Seligman's reformulated theory of learned helplessness, semantic network theories, and Beck's concept of schemata. These three models are selected because of their relevance to depression and Power critically discusses them in relation to the requirements of a theory of depression on the one hand, and on the other in relation to the requirements of an adequate cognitive model. The former includes the meaning of depression and its prevalence, the patient's history, and the contribution of contextual aspects such as loss and failure. The latter refer to the form of knowledge representation, for example, networks and schemata and the units of representation they employ, and to the mode of organization, as for example a hierarchy or heterarchy.

Power proposes that both sets of requirements can be accommodated within a mental models approach. This can subsume both a propositional level of representation and higher-level schema-type information about the self, goals, and relationships, all of which are relevant to depression. Another advantage is the way in which mental models can form a heterarchical organization such that one model will appear superordinate to another under some conditions but subordinate to that same model under other conditions. The negative model of the self may be dominant during an episode of depression and a positive one present between episodes of depression. Such a system corresponds with an ambivalent model of the self in depression.

Although a number of quite distinct theoretical traditions underlie the development of structural as compared with processing models, any contrast of this type obviously oversimplifies the state of affairs. There has been a substantial conceptual overlap and interchange between these research traditions in cognitive psychology and Eysenck and Mathews recognize that cognitive approaches within the clinical context that variously deal with stimulus content, with structure, and with ways in which information is processed, must ultimately interact with one another. Mathews and Eysenck illustrate this with reference to vulnerability to anxiety disorders. These may arise as a result of a systematic preattentive bias in cognitive input to the emotional evaluative system that results in the more threatening aspects of events being successful in capturing processing resources. Once captured, selective attention ensures the additional intake of information from sources of threat, which in turn leads to their preferential encoding in memory. Differences in long-term memory, such as can be postulated, for example, between those who are high in trait anxiety compared with those who are low, may affect processing efficiency and hence reactions to threatening stimuli. Hence a cycle is formed of inefficient processing, inappropriate reactions to threat, and selective storage in memory.

Some of the questions raised by cognitive models are of particular importance to behavior therapists. They include the role of introspective evidence and reported thought content, the role of conscious and unconscious processes, and the postulate of a central control system or executor that regulates behavior. Although the problems may be shared by theorists and clinical practitioners, the solutions are sought in very different ways.

We can consider two examples: the potential conscious access to mental states and the role of self-report. These are relevant to experimental cognitive psychologists for a variety of reasons, which include the nature and structure of cognitive functions and their permeability or encapsulation from one another; and how processes such as perceptual processes get translated into what we consciously experience. The orientation for clinically oriented psychologists is different. They have been viewed as the processes within an individual that mediate or cause certain behaviors. The behavioral-cognitive approach emphasizes the capacity of humans to use the verbal-representational system to regulate their own behavior. Verbalizable knowledge and self-awareness are seen as essential prerequisites for the control of behavior.

Brewin considers some views on the potential conscious access to mental states. One is that nonconscious cognitive material can become accessible to consciousness without much distortion of the unconscious material. This is common to dynamic psychotherapists and in part to those cognitive therapists who assume that dysfunctional beliefs of which the patient is unaware can be identified from contents of automatic thoughts, thinking errors, use of particular words, etc. Others assume a complete separation between conscious and nonconscious processes. Midway between these two extremes are those who postulate some transformational process between what is automatically processed and the individual's phenomenal experience of it.

The other issue referred to—the nature and role of verbalizations—has generated an extraordinarily diverse range of comment from the contributors to this book. Verbalizations can be automatic, (Williams' discussion of Beck's cognitive therapy, page 257) that is, automatic by virtue of their coming out of the blue, seemingly unprompted by events. They are used in self-talk therapy to provide information about thinking styles, because observers can note the content of thought, through speech, much as one can observe overt behavior. They convey affect (cf. Eifert) and affective descriptions of the self (e.g., "I'm a failure"), which can be counteracted by positive self-statements. They are discussed by Brewin, and Evans and Litz, in relation to whether and under what conditions conscious thought relates to behavior. Brewin differentiates verbal reports of behavior in terms of whether action is intended or unintended, regulated or unregulated.

They are also a means by which individuals can summarize their knowledge of the rules of the environment, exemplified in Chapter 6 in terms of information about stimulus sequences. Discrepancies between knowledge and use of knowledge are observed, which relate to the clinically relevant distinction between knowing the rules and being able to put them into practice. Lowe *et al.* suggest that the development of language in humans facilitates rule-governed behavior.

This lively interest in the use of verbal report in the clinical context again illustrates wide differences between clinical and experimental cognitive approaches.

Semantic analysis is firmly at the center of much cognitive research, yet with a totally different perspective from the clinically oriented interests described here.

Clinical concerns prompt many questions about kinds of non-conscious representations and how they exert an influence on attitudes and actions related to maladaptive behavior. There is a strong need to tackle this question, and Greenberg and Safran (Chapter 14) speak of emotional representation in terms of inferring a structure in memory that consists of episodic memories, eliciting environmental events, expressive motor response, autonomic arousal and conceptual appraisal. Their own view derives from multilevel, parallel processing of information from both within and outside the individual that is integrated outside awareness. There is growing convergence of views that the representation of emotions must include not only verbal propositional statements but, more importantly, motor-effector components. Greenberg and Safran refer to “action disposition information” as ultimately leading to action. Dysfunction can thus occur in two ways, one from the type of action disposition generated and another from the type of processing this information undergoes.

The kinds of schemata arising from clinical requirements are also likely to include psychophysiological data, and these form a component of most of the models that Greenberg and Safran discuss. Lang’s extension of network theory to include psychophysiological response elements, verbal statements, and overt behavior is an example of a model that is plausible and relevant to theory and therapy. Lang suggests that a particular emotional experience is a construction of expressive-motor reactions, visceral and somatic components, ideas and memories, and that the activation of any one of these components evokes other parts of the network. The “data structure” according to Lang is one in which emotion information is coded in memory in the form of propositions that are organized into an associative network. This prototype network is associated with a production system that is an information analysis program (i.e., the emotional image) and a program for response generation. Somato-visceral efferents and action are the output events occasioned by the production system processes of response information.

From this analysis, Lang suggests three major issues that require further examination—valence, arousal, and control—and that provide a natural classification of methods used in the treatment of anxiety disorders.

That is to say, therapists either attempt to reduce arousal through drugs or relaxation training . . . or try to modify the valence of negative contexts by for example, providing success experiences in the aversive context and by reinterpreting the meaning of negative situations . . . or they focus on training the subject in self-control, efficacy and competency, by teaching skills and modelling.

To be maximally effective, therapy may need to consider all three parameters.

This kind of theory illustrates a number of requirements that behavior therapists might expect it to cover. It allows for the role of conditioning as a mechanism by which the network develops to its current form (although the detail of how this comes about remains unspecified at present) and it provides an integrated view of cognition and emotion. The implication is that conditioning and cognitive approaches can be combined to develop representational models that, if they include response elements, are likely to be more satisfactory in the context of maladaptive behavior than prevailing affectless knowledge networks. Therapeutic intervention procedures are seen

to flow from such models (cf. Greenberg & Safran) in particular methods of affective change that can supplement the existing repertoire of behavioral and cognitive methods.

Such models violate many of the formal requirements of an adequate cognitive model. It is not clear how the idea of propositionally coded information can be extended to include response units, or how the structure of the prototype as an action set can be formalized.

What is interesting about these models is the way they differ from the laboratory-based theories that are familiar to us. These latter are concerned with a narrow range of phenomena, exactly controlled. Their strategy is to make precise predictions that can be falsified. Theories emerging from clinical practice are different in character. They are broad in scope, deal with very general kinds of real-world behavior, and have none of the detailed resolution of narrow, more exact theories. They seem to permit a smoother, more immediate translation from theory into strategies of clinical practice. They involve a conception of behavioral analysis that several of the clinically oriented contributors see as the current most important issue.

BEHAVIOR THERAPY: MODELS AND VIEWS OF BEHAVIOR

Two points emerge from the clinically oriented contributors. Theory must be relevant to clinical practice. Behavior must be defined in such a way that it is relevant to therapy. Available conditioning and cognitive models represent deep but unrelated shafts of knowledge, with little to connect them to one another or to enlighten the immediate clinical problem: how is maladaptive behavior to be treated. Their conception of behavior is confined to segmented and seemingly irrelevant behavioral units.

As a result, behavior therapists have begun to generate their own models, as illustrated in the last section. These ignore many of the distinctions between conditioning and cognition, verbal report and behavior, sensory and motor components, to give only a few examples. Data from all these sources are combined within a model that focuses directly on the relevant issue: emotionality and maladaptive behavior.

Lang's model has already been discussed. Greenberg and Safran take the integrative approach even further, by including a role for conditioning and the general biological basis of emotion. The basic structure of emotional experience is assumed to lie within a central neural program that becomes elaborated as a result of conditioning and automatic information processing and synthesizes information from various environmental sources. Subjective evaluations of significant factors of the environment activate networks and are integrated within them. Automatic appraisals occur in parallel with higher-level conceptual processes, and the information generated by a synthetic process is stored in memory as an organized structure.

These memory structures incorporate action tendencies or dispositions that can elicit goal-directed behavior when called to do so; they are mediated through a higher level of information processing, which decides which behavior is appropriate. Hence there is no fixed emotional response. Rather the emotional synthesis process generates action disposition information that is subjected to further processing to decide whether and which type of action will occur.

Such a model is quite in keeping with all the conditioning and cognitive models that have been discussed by other authors, and is an attempt to integrate many components of conditioning paradigms and theory—fear conditioning, evaluations, registration of contiguities and contingencies, psychophysiological changes, programmed activities and the language of cognition—associative networks, memory, information processing, perceptual bias, and so on. Such a general model may be lacking precision in its formal structure, but it goes some way toward building a theory that has immediate meaning and that, the authors argue, has clinical and therapeutic relevance in that it suggests intervention procedures that are behavioral, cognitive, and affective. They, like Williams, are uneasy with the rigid distinction that has been made between affective, cognitive, and behavioral systems, and question its conceptual basis. The time may not be too far off, in the clinical context, when other rigid paradigmatic distinctions such as those relating to conditioning and cognition may be reexamined.

The second issue referred to is that behavior must be defined in such a way that it is relevant to therapy. The chapters by Evans and Litz, Hallam and O'Connor, present emphatic views on this question.

Many behavioral assessment procedures are available to analyze the antecedents and consequences for each target behavior that the patient and therapist decide to pursue. One question pertains to the kind of behavioral unit or "chunk" that is relevant to therapy. The response unit that is isolated for conditioning research is not easily compounded into the larger, more meaningful behavioral descriptions utilized by clinicians in their analyses. Another question concerns the theoretical implications of behavior assessment. Evans and Litz are joined by Hallam and by O'Connor in the view that a rationale is essential, though they differ as to its nature. They are all concerned with how behavior is to be assessed and how such assessments can most fruitfully be related to effective intervention.

Evans and Litz deal directly with both issues. They chart the shifts and developments in the methodology of behavioral assessment that have taken place, and how it is coming closer to clinical needs. The original concern with measurement utilized for classification and diagnosis soon shifted to one which fastened on easily quantified aspects of behavior in order to assess the effectiveness of the intervention. This kind of behavior assessment, designed to reflect the truest measure of the phenomenon of interest, ran into many problems. An alteration in a particular parameter of target behavior might not be practically meaningful; a trivial behavioral element might be used as the outcome measure instead of the complex phenomenon itself. It means that whenever a single dimension or property of behavior is measured one runs the risk of not adequately representing phenomena that have multidimensional properties.

An alternative is to list and successively treat all problem behaviors, an approach with obvious inadequacies. The need is to organize information about complaints and to select which of a wide variety of possible intervention targets would be the most productive. This then leads on to the interesting theme of the quality of judgment and decision making that is needed for effective intervention, and the value of incorporating models of decision making into behavioral assessment.

Another difficulty is in the selection of the most problematic target for intervention. Given a particular problem, for example, aggression in a child, the target of therapeutic intervention must be judged by the therapist. It might be the lack of adaptive social skills for dealing with interpersonal conflicts, the parental reinforcement for acts of physical aggression, or the child's interpretations of other children's signals. Treatment could be directed to all three, but ideally, Evans and Litz suggest, appropriate assessment procedures should be able to ascertain which of the various intervention targets would be the most productive.

An important theme, therefore, is the quality of judgment and decision making that is needed for effective intervention. On the whole, these authors point out, there is little knowledge of whether behavioral assessment techniques protect clinicians from judgment errors or add a new set of assumptions that bias the process, for example, confirmation biases, overconfidence, and the illusion of validity. They suggest the value of incorporating models of decision making into behavioral assessment.

Hallam also argues for the functional analysis of behavior in its current environmental context, the emphasis being that concepts of meaningful action (acts) must be integrated into theories relating to behavior therapy. Content and process enter into the description of problems and the selection of targets. Hallam's case is that we cannot afford to neglect the social and historical significance of behavior, because even so-called meaningless actions, such as obsessional rituals, may be inappropriate forms of meaningful acts.

Meaningful acts are distinguished from simple reflexes and units of behavior defined in everyday terms, for example, smiling, or moving away. They imply a macroscopic level of analysis in which behavior is chunked into units to reflect a behavioral script, that is, a sequence of behavior that can be construed as an unfolding reflection of a dramatic plot. Such behavioral scripts are embodied in the real world, and are not to be confused with cognitive scripts in the mind of the individual. A behavioral approach along these lines, Hallam suggests, would make inferences about the social construction of meaning. This kind of sociological understanding is needed to supplement the natural science foundation of behavior therapy.

O'Connor reflects on the many uses of the term *behavior* and notes the lack of definitional consensus that surrounds its use. He, too, proposes a contextual approach to behavior but here context refers not to environmental stimuli or cultural norms but directly to the response process or class of processes that specify the individual's intentional relation to the world. This approach defines behavior in terms of behavioral processes alone. It requires that the context of response processes is built up empirically from independent acts: the context that defines the act is the response class of which it is a member.

The task of process analysis, differentiated from functional analysis, is to describe the context of the response that optimally accounts for the emergence of behavior. Its goal is a process dimension that lists along its axes the specific actions that typify what the individual does during various degrees of presence or absence of the problem behavior.

The proposals contained in these three behaviorally oriented chapters have radical implications for theory and practice. Each suggests ways in which behavior can be analyzed to suit clinical purpose, and each finds little guidance from established

theories. The macroscopic level of their analyses and their behavior units or chunks are far removed from conditioning or cognitive accounts of how learning or knowledge is translated into action. Each adopts its own theoretical rationale and there are proposals for distinctive modes of statistical analysis.

If they illustrate the distance between the academic theoretician and practicing clinician, they also demonstrate a keen awareness of the conceptual implications of their views. They and other contributors affirm the behavioral approach and its natural science foundations. If we have now come to accept the integration of conditioning and cognitive-behavioral approaches, how can this integration extend to encompass the kind of behavioral analysis being proposed?

The theoretical framework of the present book is clearly biological. This has been a consistent theme since the founding of behavior therapy, and made explicit in Eysenck's writings. His thesis is that learning, emotion, and individual differences can best be integrated within a biological framework. He has advocated a biosocial approach, meaning that we are concerned at all times with a paradigm of heredity–environment interaction: both play their part in observable, phenotypic behavior. Such an interactionist doctrine does not exaggerate one or the other influence but seeks to discover with precision the relative contributions of the two factors in any particular situation. New methods of biometrical genetical analysis, discussed in Chapter 18, open up possibilities of analyzing not only heritability but the total genetic and environmental architecture of abnormal behavior, including the study of the relative importance of within-family and between-family environmental factors.

A biological framework necessarily deals with the organizing mechanisms of behavior within the brain, the central nervous system, and the network of hormonal pathways. Their influences are acknowledged and examined in the chapters by Gray and Kelley. We may like to think of ourselves in terms of rational decisions, but that part of the brain that subserves rational thinking is carefully wrapped around more ancient, nonrational portions of the brain that still determine much of what we do, what we aim at, and what we seek to achieve.

Many cognitive psychologists also see their study as essentially compatible with a biological approach. They share the concept that knowledge structures are evolutionary patterns of information gathering and processing, progressively shaped in response to environmental processes (e.g., Guidano, 1984). Sayre (1986) makes explicit his biologically based view of cognitive function:

Whether human cognitive capacities are fashioned by species adaptation (i.e., they are innate) or by adaptation of individual organisms (i.e., they are learned) or by some of both, they evolved from noncognitive information processing functions. This is the basic assumption behind my approach.

The question was raised earlier about how to integrate the broad spectrum of studies that analyze behavior problems. Eysenck (1985) suggested that a certain complex division of labor is needed to cope with the very specialized tasks that each group of the scientific community performs. In the example from physics that he uses, these range from abstract theorists to intermediate theorists through phenomenological theorists who build models that are useful to the experimentalists, who then carry out applied tests. These groups comprise a continuum of proximity to

physical reality, and because their tasks are so specialized, they communicate infrequently with one another.

What holds this community together is the phenomenological theorist who tends to be more closely linked to the two extremes, the abstract theoretician at one end and the experimentalist at the other, interpreting the work of the former to the latter, and reporting back results of the latter to the former. Perhaps we should recognize such a division of labor in behavior therapy and strive to maintain the interactive flow of ideas among its participants.

“If only we can bury the hatchet of past disagreements, and concentrate on what binds us together, we may yet surprise Kuhn and achieve what he thought impossible, namely a genuine paradigm in the social sciences!” (Eysenck, 1985)

REFERENCES

- Breger, L., & McGaugh, J. (1965). Critique and reformulation of “learning theory” approaches to psychotherapy and neuroses. *Psychological Bulletin*, *63*, 338–358.
- Brewer, W. F. (1974). There is no convincing evidence for operant or classical conditioning in adult humans. In W. Weimer & D. Palermo (Eds.), *Cognition and the symbolic processes*, Hillsdale, NJ: Erlbaum.
- Eysenck, H. J. (1985). Psychotherapy to behavior therapy: A paradigm shift. In D. B. Fishman, F. Rotgers, & Cyril Franks (Eds.), *Paradigms in behavior therapy: Present and promise*. New York: Springer.
- Guidano, V. F. (1984). A constructivist outline of cognitive processes. In M. A. Reda & M. J. Mahoney (Eds.), *Cognitive psychotherapies*. Cambridge, MA: Ballinger.
- Sayre, K. M. (1986). Intentionality and information processing: An alternative model for cognitive science. *Behavioral and Brain Sciences*, *9*, 121–165.

Index

- ABC analysis, 369–370
- Action
- associative networks, 303–304
 - behavior definition, 354–355
 - decision-making and, 146
 - response-process model, 354, 367
- See also* Associative networks; Behavior
- Adaptive action tendencies, 303–304
- Adrenocorticotrophic hormone (ACTH), 407–408
- attention and, 415–416
 - individual differences and, 20–22
 - stress and, 418
- See also* Hormones
- Affect
- adaptive action tendencies, 303–304
 - associative networks, 297, 298, 301, 304–305, 306–307
 - attention and, 223
 - clinical anxiety, 219, 220, 221
 - cognition and, 248–249
 - cognitive behavior therapy, 13, 175
 - depression and, 240–241, 242
 - disorders of, 257
 - emotional synthesis process, 301–303
 - hormones and, 20
 - information processing and, 295
 - language conditioning, 169, 170–171, 173–174, 179–180
 - maladaptive, 307
 - memory and, 298–299, 303
 - schema models, 299
 - self-instructional training, 176
 - semantic network theories, 240
 - trait anxiety and, 201, 202–203
- See also* Associative networks
- Age level
- depression and, 265
 - operant conditioning and, 155–156
 - semantic conditioning, 169
- Alcoholism, 159
- Ambiguity:
- clinical anxiety and, 226
 - threat reactions and, 209
- Ambivalence, 250–251
- Animal conditioning, 57–79
- acquisition of conditioning, 62–68
 - behavioral autonomy, 61–62
 - causal relevance, 66–67
 - cognition and, 58–62
 - event correlation, 62–64
 - inhibitory conditioning, 72–73
 - instrumental conditioning, 75–77
 - learned irrelevance, 66
 - long-delay conditioning, 67–68
 - Mackintosh theory, 71–72
 - mediated conditioning, 60–61
 - occasion setting, 73–75
 - Pearce-Hall theory, 69–71
 - reinforcer reevaluation, 58–60
 - selective conditioning and blocking, 64–65
 - theories of conditioning and, 68–72
 - Wagner theory, 68–69
- Animal studies
- applicability to human, 154–155
 - attention/hormones, 414–415
 - clinical anxiety, 218
 - evaluative conditioning, 119–120
 - experimental neurosis, 81
 - fear/anxiety studies, 82–83
 - hormones, 405, 407–411
 - human psychopathology, 81–82
 - human studies and, 277
 - language and, 167
 - language conditioning and, 169
 - learned helplessness, 236
 - operant conditioning, 153
- See also* Animal conditioning; Phobic fear (primate)

- Antidepressant medication. *See* Drugs
- Anxiety
 attention and, 417
 clinical neuroscience of, 425
 heredity, 387–394
 language conditioning, 170, 172–173
 motivation and, 283
 theory and, 198–199
 trait/state anxiety distinguished, 197–198
See also Clinical anxiety; State anxiety; Trait anxiety
- Applied behavior analysis. *See* Behavior modification; Operant conditioning
- Arbitrary inference, 258
- Arousal
 constitutional approaches, 361
See also Attention
- Assessment, 331–351
 behavior therapy and, 331–332
 clinical decision-making, 344–347
 cognitive personality assessment, 339–341
 evaluative conditioning and, 127–129
 integration and, 347–348
 personality and, 341–342
 reliability issues, 333–334
 target behavior selection, 342–344
 validity, 334–336
 veridicality of measurement, 336–339
See also Tests and testing
- Associative conditioning, 403
- Associative networks, 295–311
 adaptive action tendencies, 303–304
 behavior change and, 306
 clinical implications, 306–307
 comparison of models, 301
 emotional change processes, 309–310
 emotional synthesis process, 301–303
 evaluative conditioning model, 299–301
 higher level processing, 305–306
 integrated intervention, 307–309
 integrative model, 301
 models of, 296–299
 overview of, 295–296
 patterning/somatic feedback in, 304–305
 schema models, 299
- Attention
 animal hormone studies, 414–415
 anxiety and, 417
 clinical anxiety and, 222–225, 228
 definitions, 405–406
 human hormone studies, 415–416
 information processing and, 280
 trait anxiety and, 201
- Attributional Style Questionnaire, 238
- Attributions, 236, 237, 239
- Autonomic conditioning, 137–138
- Autoshaping, 120
- Aversive conditioning
 clinical anxiety, 403
 evaluative conditioning and, 126
 hormones and, 21, 423–424
- Avoidance behavior, 284–285
- Awareness, 133–151
 alternative theoretical models, 145–146
 autonomic conditioning/stimulus awareness and, 137–138
 behavior stimulus control/mental representations integration, 146–147
 clinical conditions, 147–149
 cognition and, 147
 cognition/conditioning and, 134–143
 cognition contrasted, 279
 defined, 134
 depression, 248
 eyelid-conditioning paradigm, 138–143
 information load/salience, 144–145
 knowledge and, 135–136
 knowledge/conditioning performance and, 136–137
 knowledge effects on cognitive performance, 143–147
 language conditioning and, 187–188
- Beck Depression Inventory, 259
- Behavior
 associative networks, 300–301
 awareness and, 134
 behavior therapy and, 321–322, 326–327, 456–460
 cognition and, 14, 281
 cognitive psychology and, 145
 conditioning models of, 359–361
 definitions of, 354–355
 environment and, 379
 knowledge and, 135
 language and, 167
 language conditioning and, 180–181
 motivation and, 278–279
 repertoire of, 277
 representation, 355–357
 response process model of, 353–376
 sensory models of, 357–359
 token economies and, 157–158
See also Action
- Behavioral assessment. *See* Assessment
- Behavioral autonomy, 61–62
- Behavioral psychology, 133
- Behavior modification, 37–54
 awareness and, 134
 behavior therapy and, 37, 113

- Behavior modification (*Cont.*)
 contingency management, 158–160
 definitions and principles in, 37–40
 failures of, 154
 science and technology in, 40–45
 scientific method and, 38
 token economies, 45–50, 156–158
See also Operant conditioning
- Behavior therapy, 3–35, 315–329, 451–464
 animal analog of, 435–439
 applied science, 318–320
 assessment and, 331–332
 behavior and, 321–322, 353, 456–460
 behavior modification and, 37, 113
 cognitive behavior therapy and, 12–16
 cognitive models, 456–460
 conditioning model of neurosis, 16–19
 conditioning view, 453–456
 content of behavior, 326–327
 criticisms of, 9–12
 definitions, 37, 315
 drugs and, 433–447
 exposure as necessary condition, 23–28
 fear/anxiety neurobiology, 19–23
 hormones and, 423–425
 mental states, 322–326
 motivation and, 278
 neurosis concept and, 4–9
 phobias, 84
 psychoanalysis and, 15
 psychotherapy and, 3–4
 technical/theoretical eclecticism in, 316–318
 theory's role in, 316
- Beta-endorphin, 410–411
- Blink reflex, 138–143
- Blocking
 animal conditioning, 64–65
 long-delay conditioning, 67
- Brain, 217–218
- Brain damage, 180
- Causal relevance, 66–67
- Children
 language conditioning and, 180
 operant conditioning and, 155–156
See also Age level
- Classical conditioning
 evaluative conditioning and, 117–118
 language conditioning and, 169
 preferences, 114–117
- Clinical anxiety, 217–234
 attention and, 222–225
 behavioral therapy, 258
 cognitive content in, 221–222
 cognitive model of, 219–221
- Clinical anxiety (*Cont.*)
 cognitive research and, 228–232
 hormones and, 403–432
 incubation (in animals), 411–414
 judgment in, 226–228
 memory and, 225–226
 models of, 217–218
 phobias, 218–219
 trait anxiety and, 217
See also Anxiety; State anxiety; Trait anxiety
- Clinical assessment. *See* Assessment
- Cognition
 affect and, 248–249
 animal conditioning and, 58–62
 architecture of, 147
 associative networks, 298
 autonomic responding and, 137
 awareness and, 134–143
 behavior and, 14
 behavior therapy and, 322–326
 clinical anxiety and, 217–234
 consciousness contrasted, 279
 constitutional approaches, 362–363
 depression and, 170, 228, 235–236, 258
 evaluative conditioning and, 120
 exposure/extinction and, 27–28
 inhibitory conditioning, 72
 knowledge and, 143–147
 mediated conditioning, 61
 modification procedures for, 124
 motivation and, 277–293
 performance deficits, 200–203
 stimulus control and, 146–147
 threat reactions, 206–213
 trait anxiety and, 197–216
 unconscious and, 133–134
See also Associative networks
- Cognitive-behavior therapy
 claims of, 12–16
 depression, 257–275
 language conditioning and, 174–178
 motivation and, 278
 process research (depression), 266–268
 treatment package analysis (depression), 261–264
 treatment response predictors (depression), 265–266
- Cognitive models, 456–460
- Cognitive personality assessment, 339–341
- Cognitive psychology
 awareness and, 133
 behavior and, 145
 behavior therapy and, 317
 depression and, 268–271
- Cognitive restructuring, 177–178

- Cognitive therapy
 clinical anxiety, 231–232
 depression, 243–245
- Color-naming, 223
- Community. *See* Society
- Computer, 133
- Conditioned emotional response, 121
- Conditioned responses
 awareness and, 135
 knowledge and, 136–137
 semantic conditioning, 168–169
- Conditioned stimulus, 118
- Conditioning
 awareness and, 134–143
 behavior therapy, 453–456
 defined, 405–406
 eyelid-conditioning paradigm, 138–143
 language and, 15
 phobias, 218
See also Evaluative conditioning
- Conditioning models, 359–361
- Conditioning theory
 animal conditioning and, 68–72
 behavior therapy and, 16–19
 behavior-therapy criticism and, 10–11
 exposure and, 23–28
 neobehaviorism and, 57
 neurotic paradox and, 5
See also Animal conditioning
- Conscious awareness. *See* Awareness
- Consciousness
 unconscious and, 135
See also Awareness
- Constitutional approaches, 257, 361–363
 cognitive theories, 362–363
 physiological theories, 361–362
- Contingency management, 158–160
- Control
 eyelid-conditioning paradigm, 139
 language conditioning and, 187–188
 learned helplessness and, 236–238
- Corticosterone, 409–410
- Corticotropin, 407–408
- Corticotropin-releasing factor, 410
- Cortisol, 22, 409–410
- Decision-making
 action and, 146
 assessment and, 344–347
- Delusional speech, 177
- Depression
 Beck's cognitive therapy and, 243–245
 cognition and, 228, 245–249
 cognitive psychology and, 268–271
 cognitive restructuring and, 177–178
- Depression (*Cont.*)
 cognitive theories of, 235–236
 cognitive treatment of, 257–275
 heredity and, 387–394
 language conditioning, 170–171
 learned helplessness theory, 236–240
 life events, 219
 memory and, 225
 motivation and, 283
 process of, 251–252
 semantic networks and, 240–243
 trait anxiety and, 204
 vulnerability to, 249–251
- Depressogenic schemata, 258
- Desensitization
 clinical anxiety, 229
 depression, 258
 language conditioning, 172–173
- Diagnosis. *See* Assessment
- Diagnostic and Statistical Manual of Mental Disorders*, Third Edition (DSM-III), 5
- Dialectical behaviorism. *See* Neobehaviorism
- Dichotomous thinking, 258
- Drugs, 433–447
 clinical implications of, 439–441
 hormones, 405
 partial reinforcement/partial punishment effects, 433–439
 positive directions, 441–445
- Dysfunctional Attitude Scale, 266
- Eclecticism, 11, 316–318
- Environment
 associative networks and, 302
 behavior and, 379
 heredity and, 380–381
 response-process model, 369
- Ethics, 47–48
- Evaluative conditioning, 113–131
 antecedents of, 118–121
 associative networks, 299–301
 classical conditioning, 114–118
 clinical assessment, 127–129
 paradigm of, 115–117
 theory of, 117–124
 therapy and, 124–125
 treatment techniques and, 126–127
- Event correlation, 62–64
- Existentialism, 7
- Expectancies, 283–286
- Extinction
 exposure and, 23–28
 fear/anxiety neurobiology, 19–23
 fear of snakes (primate), 85–88
 hormones and, 407–411

- Extinction (*Cont.*)
 language conditioning, 173–174
 learning-theory criticism and, 11–12
 neurosis and, 6–7
- Extroversion–introversion
 heredity, 384–385
 neurosis and, 6, 7
- Eyelid-conditioning paradigm, 138–143
- Family, 9
- Feedback, 304–305
- Fear/anxiety reactions
 animal studies, 82–83
 cognitive behavior therapy and, 13
 conditioning theory, 16–19
 exposure as necessary condition, 23–28
 neurobiology of incubation of, 19–23
See also Anxiety
- Flight behavior, 304
- Flooding therapy
 depression, 258
 fear of snakes (primate), 85–88
- Functional approaches, 356
- Galvanic skin response, 168, 169, 173–174
- Genetics. *See* Heredity
- Goals, 286–289
- Hedonic transfer. *See* Evaluative conditioning
- Helplessness. *See* Learned helplessness
- Heredity, 379–402
 anxiety/depression/neuroticism, 387–394
 behaviorists and, 379–380
 definitions, 381
 environment and, 380–381
 extraversion/neuroticism, 384–385
 individual populations, 382
 methodologies, 381–382
 neuroticism and, 383
 personality and, 383–384
 psychiatry and, 380
 sex differences, 384–386
 trait anxiety, 198
- Hormones, 403–432
 attention/anxiety and, 414–417
 basic concepts in, 404–406
 clinical neuroscience, 425
 definitions, 406
 extinction of associations (animal studies),
 407–411
 extinction resistance and, 19–23
 historical perspective on, 404
 incubation of anxiety (in animals), 411–414
 neurosis and, 417–422
 treatment and, 423–425
- Hypochondria, 288
- Incubation
 fear/anxiety neurobiology, 19–23
 hormones and, 411–414
- Individual differences
 environmental events, 210
 hormones and, 20–21, 403
 language conditioning, 168, 169–170, 183–186
 memory, 203–204
 mood state, 203
 phobic fears (primate), 91–93
 threat reactions, 208, 209
 trait anxiety, 197
- Infancy, 155
- Information load, 144–145
- Information processing
 affect and, 295, 302
 associative networks, 297, 305–306
 awareness and, 133
 clinical anxiety, 220, 224
 cognitive systems and, 208
 depression, 243, 244
 trait anxiety, 199–200
- Inhibitory conditioning
 animal conditioning, 72–73
 occasion setting, 73
- Instrumental conditioning
 animal conditioning, 75–77
 neurosis and, 5
- Intelligence, 265
- International Classification of Diseases, 6
- Introversion–extroversion. *See* Extraversion–
 introversion
- Judgment, 226–228
- Knowledge
 awareness and, 135–136
 behavior and, 135
 cognitive performance and, 143–147
 conditioning performance and, 136–137
 reinforcement and, 134
 theoretical models, 145–146
- Language
 clinical anxiety and, 223
 conditioning and, 15
 mental models and, 246
- Language conditioning, 165–193
 affect and, 179–180
 anxiety, 170
 applications of, 171–178
 awareness and, 187–188
 behavioral changes and, 180–181
 behavior and, 167
 cognitive-behavioral interventions, 174–178

- Language conditioning (*Cont.*)
 depression, 170–171
 language function and, 169–170
 semantic/verbal conditioning interventions,
 172–174
 terms/concepts in, 168–169
 variables influencing effectiveness of, 183–188
 verbal behavior and, 178–179
- Latent inhibition effect, 66
- Law, 47–48
- Learned helplessness
 depression and, 236–240, 258–259
 reality and, 245
- Learned irrelevance, 66
- Learning theory
 animal conditioning and, 57–79
 animal/human behavior, 154–155
 assessment and, 332
 associative networks, 300
 behavior modification and, 41, 113
 behavior therapy and, 317, 318
 behavior-therapy criticism and, 9–11
 cognitive behavior therapy and, 15–16
 expectancies, 283
 fear/anxiety, 26–27
 language and, 167
 phobias and, 88–89
See also Behavior modification
- Levels and filters, 5–6
- Life events
 clinical anxiety, 219
 depression, 245, 250, 252, 258
- Locus of control, 283
- Long-delay conditioning, 67–68
- Long-term memory
 trait anxiety and, 199–200, 203–206
See also Memory; Short-term memory
- Magnification, 258
- Measurement
 assessment and, 331
 veridicality of, 336–339
See also Assessment; Tests and testing
- Mediated conditioning, 60–61
- Medical disorders, 6
- Medicine, 41
- Memory
 affect and, 298–299, 303
 associative networks, 297, 298, 302–303
 clinical anxiety and, 225–226
 depression and, 240, 269–271
 psychopharmacology, 405
See also Long-term memory; Short-term mem-
 ory; Working memory
- Mental illness, 5–6
- Mental models, 245–247
- Mental states, 280–281
- Meta-analysis
 cognitive-behavioral intervention, 175
 psychotherapy and, 3
- Minimization, 258
- Modeling, 218
- Modern neobehaviorism. *See* Neobehaviorism
- Mood
 affective disorders, 257
 associative networks, 298
 clinical anxiety and, 224
 trait anxiety and, 203
- Motivation, 277–293
 centrality of, 278
 clinical implications, 289–291
 cognitive systems and, 279–282
 expectancies, 283–286
 values/goals, 286–289
- Neobehaviorism
 behaviorism and, 15
 cognitive behavior therapy and, 12–13
 conditioning theory and, 57
- Network models. *See* Associative networks;
 Semantic network models
- Neuroendocrine system. *See* Hormones
- Neuropsychology
 emotional synthesis process, 301–302
 response process model, 363–364
- Neurosis
 anxiety, 197
 behavior therapy and, 4–9
 behavior therapy/psychotherapy compared, 3
 conditioning model of, 16–19
 definitions, 5, 405–406
 experimental, 81
 exposure and, 23–28
 extraversion-introversion, 6, 7
 heredity and, 379–402
 hormones and, 20–21, 403, 417–422
 medical disorders and, 6
 memory and, 203–204, 225–226
 mood and, 203
 Pavlovian conditioning and, 6–7
 phobic disorders and, 9
 psychosis and, 6
 society and, 5–6
 vasopressin and, 419
- Neurotic paradox
 behavior therapy and, 5
 conditioning models, 360–361
- Observational conditioning
 animal/human compared, 103–106
 phobic fears (primate), 88–104

- Obsessive-compulsive disorder, 219
- Occasion setting, 73–75
- Operant conditioning, 153–165
 adult/infant compared, 155
 age level and, 155–156
 animal/human behavior, 154–155
 animal studies and, 153
 behavior modification and, 38
 contingency management, 158–160
 failures of, 154
 token economies, 156–158
 verbal behavior, 168
See also Behavior modification
- Operant learning theory. *See* Behavior modification
- Overgeneralization, 258
- Oxytocin, 411
- Panic attacks
 clinical anxiety, 217, 218–219, 221–222
 response process model, 371–373
- Partial punishment effect
 drugs and, 435–439
 phobias, 433–435
- Partial reinforcement effect
 drugs and, 435–439
 phobias, 433–435
- Passivity, 236
- Pavlovian conditioning, 5, 6–7
- Perception, 133
- Personality
 assessment purpose and, 341–342
 cognitive-behavioral intervention, 175–176
 cognitive personality assessment, 339–341
 heredity and, 382, 383–384
 language conditioning and, 183–186
- Personalization, 258
- Phobias
 attention and, 222–223
 behavioral therapy, 258
 clinical anxiety and, 218–219
 distribution of, 95–102
 expectancies, 284–285
 heredity, 390–391, 392–393, 394
 hormones and, 20
 language conditioning and, 172–174
 neurosis and, 9
 partial reinforcement/punishment effects, 433–435
 trauma and, 88
 treatment, 229
- Phobic fears (primate), 81–111
 etiology through observational conditioning, 88–103
 immunization/prevention of fears, 103–106
- Phobic fears (*Cont.*)
 measurement of fear, 84–85
 symptomatology of fear of snakes, 83–85
 therapy for fear of snakes, 85–88
- Placebo effect
 behavior therapy/psychotherapy compared, 3
 expectancies, 284
- Positivism. *See* Scientific method
- Preferences. *See* Evaluative conditioning
- Premack principle, 261
- Preparedness hypothesis, 392–393
- Primate phobic fears. *See* Phobic fears (primate)
- Prioritization, 287
- Problem solving, 200–203
- Psychiatry, 380
- Psychoanalysis
 behavior therapy and, 15
 community and, 6
 incubation and, 24
- Psychopathology
 animal models of, 81–82
 behavior therapy and, 320
 token economies and, 156–158
- Psychopharmacology. *See* Drugs
- Psychosis, 6
- Psychotherapy
 associative networks, 306
 behavior therapy and, 3–4
 motivation and, 290
 varieties of, 257
 verbal behavior and, 178–179
- Recall. *See* Long-term memory; Memory; Short-term memory; Working memory
- Reinforcement
 adult/infant compared, 155
 animal conditioning and, 58–60
 animal/human compared, 155
 causal relevance, 66–67
 conditioning models, 359
 depression, 261
 eyelid-conditioning paradigm, 139
 inhibitory conditioning, 72–73
 knowledge and, 134
 language conditioning, 169, 171, 181
 learned helplessness, 236
 learned irrelevance, 66
 long-delay conditioning, 67–68
 mediated conditioning, 60–61
 motivation, 278
 occasion setting, 73–75
 phobias, 433–435
 semantic conditioning, 168–169
 token economies, 157–158

- Relaxation
 clinical anxiety, 229
 language conditioning, 172–172
- Reliability, 333–334
- Response process model, 353–376
 action and, 354
 behavior as representation, 355–357
 behavior as response controlled, 363–367
 behavior definitions, 354–355
 clinical example in, 371–373
 conditioning models of behavior, 359–361
 constitutional approaches to behavior, 361–363
 response-centered approach, 367–373
 sensory models of behavior, 357–359
- Salience, 144–145
- Schemata
 associative networks, 299
 depression, 243, 244, 269
 depressogenic, 258
- Schizophrenia, 177
- Scientific method
 awareness and, 134
 behavior modification and, 38, 40–45
 behavior therapy and, 7–9, 318–320, 322
 behavior-therapy criticism and, 11
 behavior therapy/psychoanalysis compared, 15
- Selective abstraction, 258
- Selective conditioning, 64–65
- Self-awareness. *See* Awareness
- Self-efficacy, 284–285
- Self-esteem
 depression and, 236, 243–244, 250
 language conditioning, 171
 verbal conditioning, 177–178
- Self-handicapping, 288
- Self-instructional training, 176–177
- Self-referent thought, 286
- Self-report
 anxiety scales, 197
 awareness and, 135–136
 cognitive therapy and, 279
 depression, 250–251, 262, 269
 information load and, 144
 mental states and, 280–282
 reliability, 334
See also Tests and testing
- Semantic conditioning
 applications of, 172–174
 behavior and, 181
 defined, 168–169
- Semantic network
 associationist models and, 296–297
- Semantic network (*Cont.*)
 depression and, 240–243
- Sensory models, 357–359
- Signal detection theory (SDT), 358
- Skin conductance responding, 392
See also Galvanic skin response
- Social constructivism, 324
- Social learning theory. *See* Learning theory
- Social networks
 behavior therapy and, 9
 depression and, 250
- Society, 5–6
- Somatic feedback, 304–305
- State anxiety
 threat reactions and, 208
 trait anxiety distinguished from, 197–198
See also Anxiety; Clinical anxiety; Trait anxiety
- Stimulus awareness, 137–138
- Stimulus control, 146–147
- Stimulus–response (S–R)
 awareness and, 134
 behavior definition and, 354
 conditioning models, 360
 constitutional approaches, 362–363
 sensory models, 357–359
- Stress
 hormones and, 418–419
 neurosis and, 403
- Subjective reactions, 13
- Suicide
 depression, 270–271
 language conditioning, 170–171
- Task performance, 200–203
- Technology, 40–45
- Tests and testing
 anxiety, 199
 depression, 259
 measurement veridicality, 336–339
 threat reactions, 207–208
See also Assessment; Self-report
- Therapeutic relationship, 183–184
- Threat reactions, 206–213
- Time-out procedures, 47–48
- Token economies
 behavior modification, 45–50
 operant conditioning, 156–158
- Trait anxiety, 197–216
 anxiety neurosis and, 198–199
 clinical anxiety and, 217
 cognitive system and, 199–200
 defined, 197
 genetic factors in, 198
 long-term memory and, 203–206

- Trait anxiety (*Cont.*)
performance deficits and, 200–203
state anxiety distinguished from, 197–198
theoretical understanding of, 198–199
threat reactions, 206–213
- Trauma
conditioning model of neurosis, 16–17
phobias and, 88
- Treatment
assessment and, 342
associative networks, 306–310
awareness and, 147–149
behavior therapy/psychotherapy compared, 3
clinical anxiety, 229–230
depression and, 240, 257–275
emotion/cognition/action network, 295–296
evaluative conditioning and, 126–127
expectancies and, 283–286
hormones and, 423–425
motivation, 289–291
phobias, 229
target behavior selection, 342–344
- Twin studies. *See* Heredity
- Unconditioned stimulus
classical conditioning, 118
hormones, 405
semantic conditioning, 168
- Unconscious
cognition and, 133–134
consciousness and, 135
depression and, 248
- Unconscious (*Cont.*)
perception and, 133
- Validity, 334–336
- Values
behavior therapy, 316
motivation, 286–289
- Vasopressin, 22, 411, 419
- Verbal behavior
contingency management, 158–160
language conditioning and, 178–179
operant conditioning, 168
token economies, 156–158
- Verbal conditioning
applications of, 172–174
defined, 168
depression, 177–178
evaluative conditioning and, 118
schizophrenia, 177
- Verbalizable knowledge. *See* Knowledge
- Verbal report. *See* Self-report; Tests and testing
- Vulnerability, 243–244, 249–251
- Weight reduction, 159–160
- Wisconsin General Test Apparatus (WGTA),
84, 89–90
- Working memory
depression, 241–242
trait anxiety, 202
See also Long-term memory; Memory; Short-term memory
- Worry, 204–205, 222, 227–228