BEHAVIOR THERAPY AND THE CONDITIONING MODEL OF NEUROSIS

Hans J. EYSENCK *
University of London, UK

The methods of applied science are usually based on the findings of pure science, and the discovery of successful methods of therapy for the neuroses is dependent on the formulation of appropriate theories for these disorders. It is argued that the only viable theory about neurosis at the moment is an adaptation of the Watsonian conditioning model, suitably altered to fit in with more recent discoveries in the fields of the formation and extinction of conditioned responses. Such a model has important relevance to behaviour therapy and the various methods for treating neuroses falling under that heading, and indeed for an explanation of the apparent effectiveness of psychotherapy, psychoanalysis, and whatever events mediate spontaneous remission.

Neurotic behaviour has often been defined as maladaptive behaviour accompanied by strong, irrelevant and persistent emotions, occurring in full awareness of the maladaptive and irrational nature of the behaviour in question. Typical instances of neurotic behaviour so defined can be found in anxiety reactions, phobias, obsessive-compulsive behaviour, reactive depressions, and psychosomatic symptoms. The typical neurotic reactions here mentioned form a general syndrome which the writer has called "dysthymic"; it is differentiated from another group of symptoms, including hysteria, psychopathy, and antisocial behaviour generally (Eysenck and Rachman 1965). We have called these two groups of symptoms "disorders of the first and second kind", respectively; in this paper we will deal mainly with disorders of the first kind. My theory of disorders of the second kind has been

* Author's address: Hans J. Eysenck, University of London, Inst. of Psychiatry, De Crespignuy Park, Denmark Hill, London SE5 6AF, UK.

0020-7594/81/0000-0000/$02.75 © 1981 North-Holland
developed and discussed elsewhere (Eysenck 1977, 1979; Eysenck and Eysenck 1978).

It has for a long time been clear that neurotic behaviour of the kind described above presents a problem both for common sense and also for psychological theory. Mowrer (1950) has referred to this problem as the "neurotic paradox". The paradox derives from the fact that most philosophical and psychological theories are essentially hedonistic, stressing what Thorndike has called the "law of effect" and Skinner the "law of reinforcement". How, in these terms, can we account for the existence of a class of behaviours which are at the same time "self-perpetuating and self-defeating?" (Mowrer 1950: 486.)

As Mowrer states: "Common sense holds that a normal sensible man, or even a beast to the limits of his intelligence, will weigh in balance the consequences of his acts: if the net effect is favourable, the action producing it will be perpetuated; and if the net effect is unfavourable, the action producing it will be inhibited, abandoned. In neurosis, however, one sees actions which have predominantly unfavourable consequences, yet they persist over a period of months, year, or lifetime" (1950: 486).

The existence of this paradox casts doubt on motivational theories otherwise well supported, and poses a problem for psychology which cannot be avoided. From the practical side, too, it is clear that we need an answer to the question of how neuroses can originate and exist, in order to make it possible for us to construct theories of treatment which would lead to better methods than those currently in use, such as psychoanalysis and psychotherapy, which have been found largely wanting (Rachman and Wilson 1981).

Most existing models of neurosis are based on observations of the behaviour of human beings in the clinical therapy situation, and interpretations of their actions and words in that situation; these interpretations are often very far-fetched, and dominated by theory (Gossop 1981). As many philosophers of science have pointed out, these models, particularly the psychoanalytic, are not put in a form which enables them to be tested experimentally and falsified; consequently they cannot be accepted as scientific theory in the usual sense of that term. The only theory which has grown out of the background of experimental and laboratory studies and which leads to verifiable consequences, is the conditioning theory of Pavlov and Watson, and it is this theory which we will be concerned with in this paper. What the theory says, in
effect, is that all neurotic disorders are produced by conditioned emotional (autonomic) responses, and behaviours instigated in the patient by these conditioned anxieties in an attempt to escape from the anxiety so produced. Thus for instance the handwashing behaviour of many obsessive-compulsive patients is explained as an attempt to reduce the anxiety produced by fear of contamination; this in turn is regarded as a conditioned response acquired at an earlier period. A great deal, of course, is known about the process of conditioning, both from animal and from human work, and consequently it is possible to produce quite strict tests of the theory in question.

This has often been denied by critics like Locke (1971), London (1972), and Breger and McGaugh (1965), who have criticised behaviour therapy for not living up to its pretentions. Eysenck (1959), in his original contribution inaugurating the term "behaviour therapy" in its current meaning, suggested that it was differentiated from all other types of therapy by its reliance on a solid, experimentally demonstrable and laboratory-based theory, namely that of conditioning and extinction, and that it consequently had a scientific status altogether lacking in the various types of psychotherapy. A discussion of some of the issues raised by the critics has been given by Eysenck (1976), with the conclusion that while there were many anomalies, as in all scientific systems, nevertheless there was a close connection between the experimental and theoretical work of classical academic psychology, and the theory of conditioning and practical application of behaviour therapy.

As an example, and to illustrate the difference between typical psychiatric and psychoanalytic "interpretative" theories, and the conditioning model, let us look at a relatively simple kind of neurotic reaction, namely enuresis. Psychoanalysts regard enuresis as a symptom of a deeper underlying disorder. According to this point of view, the clinician attaches fundamental causal importance to the deep-seated patterns of the child—parent relationships which are causally related to future neurotic behaviour. Some of the specific theories embraced by analysts take the form of highly speculative interpretations based on psychoanalytical symbolism. For one analyst, for instance, enuresis "represented a cooling of the penis, the fire of which was condemned by the super-ego". For another, enuresis was "an attempt to escape a masochistic situation, to expel outwards the destructive tendencies: the urine is seen as a corrosive fluid and the penis as a dangerous weapon". Yet another therapist suggested that usually enuresis expresses
a demand for love, and might be a form of "weeping through the bladder".

There are too many different speculations of this kind among psychoanalytic writers to list them all; they can conveniently be grouped under three different headings. Some believe that enuresis is a substitute form of gratification of repressed genital sexuality — if I can't sleep with my mother, then I'll use my penis this way. Others regard enuresis as the direct manifestation of deep-seated anxieties and fears. Yet others interpret it as a disguised form of hostility towards parents or parent substitutes which the victim does not dare to express openly — if I can't attack you openly because you are stronger, then I'll annoy you this way! All these theories insist on the primacy of some psychological "complex" and the secondary nature of the "symptom"; concern is with the former, not the latter. Consequently, treatment is long drawn out, involves searching examination of the patient's unconscious through dream interpretation, word association, and other complex methods, and enters into consideration of many aspects of the child's personality apparently irrelevant to the simple act of bedwetting. The outcome of all this complex machinery, however, does not seem commensurate with the trouble involved; there is no evidence to show that psychoanalytic treatment of children who wet their beds produces effects in excess of simple spontaneous remission, which is very frequent.

The alternative view, derived from the theory of conditioning, is that in the majority of cases enuresis may be regarded simply as a failure to acquire a habit. This "habit deficiency" is due to faulty habit training of some kind. Ordinary continence training teaches the child to respond to bladder stimulation by awakening. The child thus learns to substitute going to the toilet (or using his pot) for bedwetting; when this learning fails, enuresis is the result. A thorough investigation has shown that, although there is sometimes something physically wrong with the urinary system, bedwetting is a habit condition in nine cases out of ten. A somewhat different view is taken by those who believe that simple enuresis is a continuation into childhood of the automatic bladder reflexes of infancy, whereas in the case of more complicated types of enuresis, the child has acquired habits of urinating during sleep in response to specific environmental conditions. Both types of enuresis are ascribed to faulty training procedures (Grosse 1980).

The bell-and-blanket method was designed to produce a connection
between the conditioned stimulus (the enlarged bladder) and the desired conditioned response, *i.e.* waking up and going to the toilet. Thus the habit deficiency is cured by supplementing the usual training procedures in this specific manner. Fig. 1 shows the effectiveness of this procedure; there are three groups of equally incontinent children, one of which is treated by means of placebo treatment, which can be seen to have no effect at all over a period of six weeks. Two groups of children are treated by means of either continuous or intermittent reinforcement; it will be seen that both achieve almost perfect continence by the end of the sixth week. This experiment (Finlay *et al.* 1973) is typical of many others; the efficacy of the procedure has been such as to replace practically all other methods of treatment.

The success of the treatment should not blind one to a realization that the validity of the conditioning theory cannot be proved by the effectiveness of the treatment. Something more specific is required by way of prediction to make it possible to accept the conditioning theory as likely to be correct. One such prediction would be that there should

![Fig. 1. Mean number of wettings per week across the 6-week treatment period for continuous, intermittent, and placebo reinforcement groups. After Finlay *et al.* 1973.](image-url)
be many relapses. It is well known that the acquisition of a conditioned response is often followed by extinction when no further reinforcement is administered, and indeed it is usually reported that the bell-and-blanket method, while initially successful, is in quite a number of cases followed by relapse. Can conditioning theory help in providing suggestions for ways of avoiding the therapeutically undesirable relapses? Several have been suggested and tested, the most interesting being intermittent reinforcement. It is well known from work in the laboratory that extinction of a conditioned response takes place much more frequently after continuous than after intermittent reinforcement, and if we equate a relapse in the therapeutic situation with extinction in the conditioning situation, this being an intrinsic part of the general theory, then we would expect that in the experiment diagrammatically illustrated in fig. 1 the group receiving intermittent reinforcement (i.e. where the blanket was connected with a battery only two out of three occasions, rather than being wired up with a battery on every occasion, as in the continuous group), then this intermittent group should show significantly fewer relapses than the continuous group. This indeed was the outcome of the experiment, thus showing the value and relevance of experimental laboratory studies for the clinical treatment of a particular neurotic condition. From the theoretical point of view, the experiment quoted above is important in demonstrating that the general conditioning theory on which the bell-and-blanket method is founded can be tested by means of quite specific deductions which can be verified in the clinical situation.

Enuresis is a rather simple and clear-cut example; let us consider a rather more complex in which the animal laboratory analogue has suggested a method of treatment for a particularly difficult and intractable disorder. The disorder in question is obsessive-compulsive neurosis, and the intractability of the disorder is clearly illustrated by a quotation from Malan (1979), one of the leading British psychoanalysts. Having discussed the Freudian theory regarding obessional anxiety, he states that “Even though the psychopathology in such cases appears perfectly intelligible, accumulated practical experience suggests that often the symptom itself develops an autonomy, and no matter how extensively the pathology is interpreted and apparently worked through, the symptom remains untouched. It is apparently true, for instance, that there is no known authenticated case of an obsessional hand-washer being cured by psychoanalytic treatment. In my view, therefore, the treatment of
choice immediately becomes *behaviour therapy*” (1979: 218–219).

This view of the intractability of the disorder is confirmed by the present writer’s analysis of hundreds of case records on the files of the Maudsley Hospital. Many different treatments were tried, in addition to psychoanalysis, such as psychotherapy, lobotomy, electro-shock, drug treatment, etc., but with results which were no the whole no better than those reported for psychoanalysis by Malan.

It seems clear that mentalistic, interpretative, pharmacological and surgical procedures do not affect this disorder very much. What has the psychological laboratory to offer? The search for an animal analogue, as suggested by Eysenck and Rachman (1965), leads us to a series of studies of dogs in a shuttle-box (Solomon et al. 1953). Briefly, what these authors did was to divide a room or “box” into two parts, separated by a hurdle which could easily be jumped by the dogs. The floor on either part of the room could be electrified separately, thus giving a shock to the paws of the dogs. A flickering light served as a conditioned stimulus, and the dogs were conditioned to jump to the light by giving the dog a shock a short time after the onset of the conditioned stimulus, whichever part of the room he happened to be in. The dogs soon learned to jump in order to escape the shock, but after some time the conditioned stimulus was activated again, and the dog received another shock, finally making it jump the hurdle. After a while the dogs became conditioned and jumped from one part of the room to the other shortly after the conditioned stimulus was activated.

This conditioned response of jumping was very firmly established, and even after the electricity supply was disconnected, so that under no conditions would the dogs be shocked again, nevertheless they continued jumping for hundreds or even thousands of trials. Different methods of trying to cure them of this “neurosis” were tried, but only one was unequivocally successful. This was a method of “flooding” with response prevention. In this method the hurdle is raised so that it is too high for the dog to jump; the conditioned stimulus is then activated, producing strong fear reactions in the dog. However, where no shock is forthcoming the dog soon quietens down. A few repetitions of this procedure serve to cure the dog completely, so that he will not jump to the conditioned stimulus even though the height of the hurdle is reduced to make it possible for him to jump from one part of the room to the other.

Formally, there is a clear similarity or analogy between the obsessive-
compulsive patient who washes his hands in order to reduce anxiety due to "contamination", and the dog that jumps in order to reduce anxiety produced by the conditioned stimulus. The important question from the point of view of clinical psychology, of course, is whether this analogy is in fact an identity, and more than a verbal similarity between two essentially unconnected types of behaviour. The work of Rachman and Hodgson (1980) demonstrates very clearly that the analogy is in fact a very fruitful one, and that treatment methods for human patients can be based on it, with considerable success.

Clinical work done at the Maudsley tried to replicate as closely as possible the Solomon et al. method of treatment. Briefly, what was done was to explain to the patient in some detail the method of treatment to be adopted, and to obtain his informed consent. Thereafter he was introduced into a fairly bare room, containing little but a table and two chairs. On the table was positioned an urn, filled with all sorts of dirt and rubbish. The experimenter-therapist would plunge his hands into this dirt and ask the patient to do likewise; thereafter he was required to sit in his chair quietly, with the dirt on his hands, and refrain from going off to wash his hand, as he would, of course, have wished to do. Thus he was prevented from indulging in an activity which would have reduced anxiety, just as the dogs were prevented from the jumping activity which would have reduced their anxiety. Just like the dogs, the patients too were "flooded" with emotions, and showed strong fear and anxiety reactions. However, as in the case of the dogs, this anxiety died away over time, until after an hour or so very little remained. Further repetitions of this procedure are required, but the outcome was outstandingly successful. Something between 85% and 95% of all patients were cured or very much improved, after a relatively short period of treatment; follow-ups showed that instead of relapses or symptom substitution, patients actually improved still further in their work adjustment, sexual adjustment, etc. Details are given in the book by Rachman and Hodgson (1980); here we would merely wish to draw attention to the contrast between the complete failure of psychoanalytic methods in treating this disorder, and the success of the experiment-based method adopted here.

The hypotheses giving rise to the treatment were investigated directly, by means of patients' reports, psychophysiological studies of their autonomic reactions, etc. Fig. 2 shows mean ratings for urge to go and wash and discomfort experienced, across occasions. The measurement occa-
sions plotted on the horizontal axis are: BE, before exposure to provoking stimulus (i.e. contamination with dirt); AE, after exposure; AR, after ritual (i.e. after handwashing); and again AE, after second exposure, and half-hourly intervals up to three hours of response prevention. Thus the left side of the figure illustrates the anxiety-relieving function of the ritual (handwashing), which reduces the urge and discomfort experienced to the pre-contamination level. Similarly, on the right, it will be seen that while there is a strong outburst of emotion ("flooding") after exposure, the simple passage of time leads to extinction. A three-hour extinction period is plotted, but even after one hour extinction is almost complete. Plotted here are verbal reports, but psychophysiological measures also taken gave similar results. The theoretical conception based on experimental work with animals in the laboratory thus enables us to make predictions and generate methods of treatment.

![Fig. 2. Mean ratings for urge and discomfort across occasions. The measurement occasions plotted on the horizontal axis are: BE, before exposure to provoking stimulus; AE, after exposure; AR, after ritual; AE, after second exposure; and half hourly intervals up to three hours. From Rachman and Hodgson 1980.](image-url)
for humans which can be shown to be verified when tried out in actual practice. It will now be clear why Malan, a professional psychoanalyst, advocated behaviour therapy as a method of choice for this particular disorder!

A third and last example will be given to illustrate the application of general principles derived from laboratory work to the design of treatments for clinical patients, in this case neurotic depressions. The work was carried out by McLean and Hakstian (1979), and was concerned with a comparison between psychotherapy, behaviour therapy, drug therapy, and a placebo-type relaxation therapy. Ten different measures were used to assess the success of the treatment, and great care was taken to objectify and quantify these estimates. One hundred and ninety-six depressed patients who met the screening and selection criteria were assigned to one of the four treatment conditions. A particular feature of the study was the investigation of drop-outs from different methods of treatment; these were noted and replaced in order to keep identical numbers in the four groups. In essence, the behavior therapy employed was based on the Skinnerian hypothesis that depression is the result of lack of positive reinforcement, and that improvements in the patients' coping behaviour (communication, behavioural productivity, social interaction, assertiveness, decision making, and problem solving) would lead to greater numbers of positive reinforcements occurring, and hence to a reduction in depression. Treatment consisted essentially in increasing the gainful interaction with the environment of the patients, despite the temporary experience of depressed mood, in a manner that would lead to more frequent positive personal and social recognition. Clients were required to engage in daily skill development activities and to monitor their achievement by means of structured log sheets.

Results of the experiment showed that there was only a 5% drop out rate for the behavioural therapy group, which was significantly lower than the 30% for the psychotherapy group, and the 36% for the drug therapy group.

On practically all of the ten measures used to evaluate the effects of the therapy, the behaviour therapy group did best. On six of the ten measures, the psychotherapy group scored most poorly. As the authors conclude, “the results show unequivocal superiority for the behavioural intervention as indicated by immediate treatment response (best on 9 out of 10 outcome measures) and a more marginal superiority on fol-
low-up (best on 7 of 10 outcome measures) . . . Behaviour therapy clients were significantly superior to drug therapy clients no only on the social measure but also on the mood and complaint measures. Further, the drug therapy group did marginally better than the psychotherapy on the social outcome measure”.

Interesting in its own right is the relative success of the relaxation therapy group, which was originally included as a kind of placebo. This indicated not only that drug therapy and psychotherapy are less powerful than might have been thought, but also that a large number of non-specific variables undoubtedly influence treatment outcome. “It may be argued that relaxation therapy represents a coping technique and in this sense is an active treatment rather than a control condition. This is perhaps the case, but there is, on the other hand, no compelling theoretical argument that suggests that muscle relaxation alone is a sufficient treatment for clinical depression. According, the treatment results in this group are considered to be attributable to non-specific (i.e. placebo) effects.”

“The psychotherapy treatment proved to be least effective at both post-treatment and follow-up evaluation periods, and, generally speaking, fared worse than the treatment control condition. Considering that 30% of the clients in this group remained in the moderate-severe range of depression by the end of treatment, compared with 19% of the control condition clients, this treatment cannot be assumed to be benign. It may be helpful to note again that the therapists in the psychotherapy group were both experienced at working with their preferred treatment.”

Clearly the outcome of this study again emphasizes the success of laboratory experiment-based methods, and the utter failure of psychotherapeutic methods based on hypothetical and non-experimental conceptions and interpretations. The examples given only illustrate these points; the book by Rachman and Wilson (1980) should be consulted by readers doubting that these conclusions can be generalized to the whole field of clinical treatment.

While it is indisputable that in this field, as in science generally, successful practice can only be built upon firm scientific knowledge acquired in the laboratory, it is not often realised that the process is a reciprocal one, and that a feedback principle is operating by means of which information from the clinic may reach back to the theoretical and experimental academic side of psychology, and cause changes in
established theories because of a failure of observed events to follow prediction. A chain of events illustrating this line of argument may also be illustrated by reference to our own work on the conditioning theory of neurosis, i.e. by now looking at the origins of neurotic disorders rather than their treatment. As already remarked earlier in this chapter, the original theory of neurosis, using conditioning principles, was elaborated by Pavlov and Watson, but although it aroused much interest there were obviously many difficulties in the way of accepting it in its simple fashion, i.e. to regard all neurotic disorders and symptoms as simply conditioned autonomic responses, or anxiety-reducing behaviours instigated by the conditioned autonomic response mechanisms. The difficulties arising from this original model have been discussed in great detail by Eysenck (1979), and it is these difficulties that gave rise to certain reformulations of widely accepted laws, such as the law of extinction. Thus this process of reciprocal innovation will serve admirably to illustrate our point.

However plausible the conditioning theory of neurosis may appear, it has certain obvious weaknesses. These may be briefly listed.

(1) Watson's theory was based essentially on a single case, namely the induction of a phobic fear of rats in an 11-month old infant, the famous "little Albert" (Watson and Rayner 1920). Later investigators have been unable to replicate Watson's results (English 1929; Bregman 1934). Watson seems to have had some premonition of the likelihood that not all infants might be capable of replicating the phenomenon; as he says in his paper with Rayner: "One may possibly have to believe that such persistence of early conditioned responses will be found only in persons who are constitutionally inferior" (1920: 14). Individual differences have indeed been found to be prominent in predisposing people to neurotic disorders (Eysenck and Rachman 1965), but the notion of "constitutional inferiority" has no experimental backing or theoretical meaning, is untestable in its present form, and if taken seriously would suggest that Watson's theory is only applicable to a very small group of persons.

(2) It has often been pointed out that phobias are relatively restricted to a small set of stimuli (Geer 1965; Landy and Gaupp 1971; Lawlis 1971; Rubin et al. 1968; Wolpe and Lang 1964). This seems to contradict the notion of equipotentiality, which is implicitly accepted by Watson and his followers, and which assumes that all
stimuli are equally capable of acting as CS. As Seligman (1971: 312) has pointed out, speaking of phobias: "They comprise a relatively nonarbitrary and limited sets of objects; agoraphobia, fear of specific animals, insect phobias, fear of heights, and fear of the dark, and so forth. All these are relatively common phobias. And only rarely, if ever, do we have pyjama phobias, grass phobias, electric-outlet phobias, hammer phobias, even though these things are likely to be associated with trauma in our world". Thus a set of potentially phobic stimuli seems to be nonarbitrary, and to be related to the survival of the human species through the long course of evolution, rather than to recent discoveries and inventions which are potentially far more rational sources of phobic fears, such as motor cars, aeroplanes, and guns. The nonarbitrary and limited choice of objects and situations which predominantly produce phobic fears in humans is difficult to explain along traditional lines, and Watson's theory therefore seems to break down in relation to this well-documented phenomenon.

(3) Single trial conditioning is assumed in Watson's theory, which relies essentially on traumatic events producing conditioned fears. However, single trial conditioning is very rare in the laboratory (Kamin 1969; Seligman 1968), and histories of the development of peace-time neurotic disorders do not usually include reference to traumatic events of any kind. Conditions are different in wartime, when traumatic events do happen quite frequently, and sometimes produce neurotic disorders of a lasting kind. The absence of such traumatic events in peace-time neurotic disorders, must throw a grave doubt on Watson's theory in its present form.

(4) It is well known that the usual CS-UCS connection is very dependent on precise experimental conditions, particularly the time relations involved. Considering eyeblink conditioning, for example, it is known that such conditioned responses can only be established when the CS precedes the UCS by between 500 msec and 2,500 msec. But such precision in unobtainable (except by chance, and occasionally) in real-life situations, and it becomes difficult, therefore, to transfer laboratory conditioning theories to everyday life, where contingencies are not controlled, and where consequently much longer periods may elapse between CS and UCS.

These and other considerations have led Seligman (1968, 1971) to formulate the hypothesis of "preparedness". According to this theory,
"phobias are highly prepared to be learned by humans, and, like other
highly prepared relationships, they are selective and resistant to extinc-
tion, learned even with degraded input, and probably non-cognitive"
(Seligman 1971: 312). What Seligman is saying, essentially, is that
through a process of evolution certain stimuli are genetically more
likely than others to become conditioned stimuli for anxiety and fear
reactions. Taking his hypothesis a stage further, it seems feasible to
postulate some kind of "instinct" theory which would suggest that cer-
tain stimuli are innately fear-producing for some people, while for
others they are "prepared" in Seligman's sense (i.e. are very easily
paired with fear-producing UCS to evoke conditioned anxiety). Presum-
ably such stimuli form a continuum, from being capable of producing
fear responses without prior conditioning, through "preparedness"
to develop conditioned fear responses readily, to what might be called
"normality", a condition in which equipotentiality reigns and the
stimuli are no more likely to produce conditioned responses than are
other types of stimuli. Presumably the position at present occupied on
the scale by a given person is in part determined by his emotional stab-
ility, and his "conditionability", i.e. his tendency to develop condi-
tioned responses quickly, strongly and lastingly.

The concept of preparedness helps to get over the above criticisms
made of Watson's theory. Thus the psychologists who failed to replicate
Watson's experiment with little Albert used common household goods,
such as curtains and blocks of wood, or a wooden duck as CSs, none of
which would have the "preparedness" value of furry animals, as used by
Watson. Similarly, lack of "equipotentiality", and the non-arbitrary and
narrow range of phobic stimuli find an explanation in terms of pre-
paredness.

Phobic fears may thus be acquired in single trials; as long as "pre-
pared" stimuli are concerned even a slight evocation of fear may be suf-
ficient to produce these already "wired-in" fear responses. Equally, the
concept of preparedness explains the fact that in typical laboratory
conditions CS-UCS intervals are of such critical importance, while in
real-life situations no such precise timing can be guaranteed. Condi-
tioning with a prepared CS can occur even with severely degraded
input, that is to say, in circumstances which under typical laboratory
conditions would lead to complete failure. Even with rats, Garcia et
al. (1971) found it possible to delay the UCS by as much as one whole
hour after presentation of the "prepared" CS, and nevertheless obtain
significant evidence of conditioning. It is thus clear that the concept of preparedness is an essential feature of any revised theory of conditioning.

It is sometimes objected that the concept is adduced specifically to get over difficulties in the conditioning theory of neurosis, and has no independent support. This is not true. There is considerable experimental evidence to demonstrate the importance of preparedness in laboratory conditioning situations too, as shown by the recent work of Hugdahl et al. (1977), and Öhman et al. (1975b, 1976). These authors have demonstrated quite clearly the absence of equipotentiality in conditioned stimuli used in their experimental studies, and the presence of "preparedness" in certain stimuli selected to test Seligman's theory.

There is another set of objections to the Watsonian model (Eysenck 1976) which cannot be explained away on the basis of "preparedness". These will now be listed.

(1) It is well known that unreinforced conditioned reactions extinguish quickly (Kimble 1961), and neurotic reactions should be no exceptions to this rule. Kimmel (1975) has pointed out that on the Watsonian model we would expect, not the development of a long-lasting neurotic condition, but rather the quick extinction of the conditioned emotional reactions. The reason of course is that the neurotic constantly encounters the fear-producing stimulus in his everyday life, or in imagination, or in his dreams, without also encountering the unconditioned stimulus. A person with a cat phobia, who may have acquired this through some form of conditioning fear producing stimuli, will in his life meet and see many cats without a replication of the unconditioned fear producing stimuli; consequently according to the traditional laws of conditioning, extinction should take place. Little Albert seeing the rats without the unconditioned stimulus of a loud noise should have extinguished his fear. As it stands this is a completely devastating criticism of the Watsonian model, and although several writers have attempted to account for the failure of extinction to occur in avoidance responding (Ritchie 1951; Solomon and Wynne 1954; Miller 1963; Soltysik 1975b), none of these seem to apply too well to the actual clinical situation, although they may have value in accounting for certain features of the laboratory animal work.

Other alternative hypotheses are reviewed by Eysenck (1979),
but even if these hypotheses could be accepted as explaining the absence of extinction, they would still fail to explain our next criticism.

(2) If the above difficulty with the Watson theory is posed by experimental laboratory studies of extinction, the next difficulty arises from clinical studies of the development of neurotic disorders. In these, we not only fail to observe the expected extinction of the unreinforced CS, but we find an incremental (enhancement) effect, such that the unreinforced CS actually produces more and more anxiety (CR) with each presentation of the CS. This fact is obvious when we consider the notion of "subtraumatic UCSs" which is sometimes introduced to salvage the Watson theory from the failure to discover traumatic UCSs. In the theory of Pavlovian conditioning, there is no provision for CRs to achieve greater strength than UCRs (Mackintosh 1974). The notion of subtraumatic UCS implies that the final CR is stronger (involves more anxiety) than the UCR. This goes counter to all we know about the fate of UCRs; these are known to habituate, rather than to increase in strength.

(3) The absence of a traumatic UCS has already been referred to briefly in connection with the fact that single trial conditioning is rare in the development of neurotic disorders. Even though in wartime, as already pointed out, traumatic events are relatively frequent in the development of neurotic disorders (Grinker and Spiegel 1945), even there many more neurotic breakdowns occur through separation from the family than through enemy action, and in peacetime neurosis traumatic UCSs are distinctly rare (Lautsch 1971; Gourney and O'Connor 1971). In the majority of cases there is some sort of insidious onset, without any single event that could be called "traumatic" even by lenient standards (Rachman 1968; Marks 1969). Thus clearly this development of neurotic illnesses is quite different from that which Watson's theory would lead us to expect, from its instigation through its later development.

(4) One last point must be made before we turn to the changes in conditioning theory which are required to make the Watson model viable. The model stresses the importance of pain in connection with the UCR, using the term to refer to simple physical pain, such as that experienced after the administration of shock, for instance. Shock, and other obviously painful stimuli, clearly mark the events in question as traumatic; the absence of traumatic events of this
kind in the development of most neuroses must cause one to doubt the omnipresence of "pain", at least in this obvious sense of the term. Watson, in his original formulation, postulated several natural causes of fear, such as loud noises, loss of support, and physical constraint; these are all "painful" in a physical sense.

From the point of view of developing a meaningful conditioning theory of neuroses, we must look for alternatives to simple pain as constituting the UCS. Thus Gray (1971) has shown that frustration ("frustrative non-reward") can have behavioural and physiological consequences identical with those of physical pain. Kimmel (1975) has suggested "uncertainty" as the basic UCS in the development of anxiety; this is usually coupled with some unpleasant stimulus whose occurrence is uncertain and therefore evades proper control, but such a stimulus need not be accompanied by physical pain (Mineka and Kihlstrom 1978). Conflict is another UCS frequently adduced theoretically in lieu of physical pain (Yates 1962). Frustration, uncertainty, uncontrollability, and conflict are of course all related, although not synonymous; they all share the characteristic of representing "mental pain" (if this term be allowed) and thus introduce a cognitive element into the conditioning paradigm. This does not make the theory a cognitive one, however; its major agent is still the simple Pavlovian conditioning and extinction process.

We have now listed a second set of four major objections to the Watsonian theory; there is no way in which these can be answered by reference to traditional learning theory, and Eysenck (1976, 1979) has suggested that what is required is a major reformulation of the law of extinction. The classical law simply states that CS not followed by UCS will lead to the extinction of the conditioned response, but Razran (1956) in his review of the state of the art at that time, already stated that "extinction continues to be clearly a less than 100% phenomenon. Instances of difficult and even impossible extinction are constantly reported by classical CR experimenters" (1956: 39).

Eysenck (1968) has suggested a brief re-statement of the law of extinction, somewhat along the following lines. "Presentation of the CS-only results in the extinction of conditioned responses of the Pavlovian A type. Presentation of the CS-only may result in extinction or incubation (enhancement) of the CR under conditions of Pavlovian B type conditioning, depending on such factors as the strength of the
UCS, the personality of the subject, and the duration of CS-only exposure.” As the differentiation between Pavlovian A and Pavlovian B conditioning is crucial to this brief statement, a few words may be apposite regarding the meaning of these terms.

Grant (1964) classified conditioning paradigms into different groupings; here we are concerned with his distinction between Pavlovian A and Pavlovian B conditioning. Pavlovian A conditioning is exemplified by the usual bell-salivation type of experiment, but Pavlovian B conditioning is different from this in many ways. As he points out: “This subclass of classical conditioning could well be called Watsonian conditioning after the Watson and Rayner (1920) experiments conditioning fear responses in Albert, but Pavlov has priority. The reference experiment for Pavlovian B conditioning might be that in which an animal is given repeated injections of morphine. The UCR to morphine involves severe nausea, profuse secretion of saliva, vomiting and then profound sleep. After repeated daily injections Pavlov’s dogs would show severe nausea and profuse secretion of saliva at the first touch of the experimenter (Pavlov 1927: 46–36). In Pavlovian B conditioning, stimulation by the UCS is not contingent on S’s instrumental acts, and hence there is less dependence upon the motivational state of the organism, and the CS appears to act as a partial substitute for the UCS. Furthermore the UCS elicits a complete UCR in Pavlovian B conditioning whereas in Pavlovian A conditioning the organism emits the UCR of approaching and ingesting the food. A great deal of interoceptive conditioning (Bykov 1957) and autonomic conditioning (Kimble 1961) apparently follows the Pavlovian B paradigm.” (See also Kalat and Rozin 1973.)

Note two major points about this differentiation of Pavlovian B from Pavlovian A conditioning. In the first place “the CS appears to act as a partial substitute for the UCS”; in other words, CS and UCS are partly, and maybe wholly, identical. Secondly, “the UCS elicits the complete UCR”; In other words the UCS provides the drive (motivation) on which the conditioning process is based. This is quite different from Pavlovian A conditioning where the dog already has to be in a state of hunger drive for the conditioning process to work at all; the conditioning process does not provide the drive.

Applied to humans, we may say that anxiety has been shown very firmly to act as a drive, so that UCSs involving anxiety fulfill the second requirement of Pavlovian B conditioning; the drive is not present prior
to the conditioning experiment, but is produced by the UCS. It will also be clear that the first condition is met; the UCR and the CR are similar if not identical, in that both involve fear/anxiety responses. Grant of course recognised this by referring to Pavlovian B conditioning in terms of the Watson-Rayner experiment; he did not go on, however, to elaborate how the distinction between Pavlovian A and B conditioning could lead to a rephrasing of the law of extinction.

What is suggested is that in Pavlovian B conditioning, CS-only presentations are not really lacking in reinforcement, because the CR which follows the CS-only is similar to, or identical with the UCR; in other words, the CR provides reinforcement for the CS-CR connection. This would produce a positive feed-back situation in which each time the CS-only was evoked, there would be reinforcement, leading to augmentation (incubation). This can occur only in situations where the UCS, and consequently the CS, provide the drive/motivation; it could not occur in Pavlovian A conditioning. This theory therefore predicts that in Pavlovian B conditioning, extinction may not occur after CS-only presentation, and that instead we may obtain incubation or enhancement of the conditioned response. The theory is developed much more fully in Eysenck (1979), and we will not go into any further details here.

Fig. 3. illustrates the point made by reference to an experiment originally reported by Napalkov in a study of dogs (see Eysenck 1976). The UCR in this experiment consisted of a blank pistol shot fired off behind the ear of the dog; the response measured was the increase in blood pressure in millimetres. The diagram presents in visual form the major results of the experiment reported by Napalkov. It shows the habituation of the UCR; when the pistol is fired a number of times behind the ear of the dog, the originally rather slight increase in blood pressure is replaced, after 25 repetitions, by a complete failure to react. Note that the response of the CR is quite different to this. After a single conditioning experience the UCS was never repeated for the dogs, nevertheless, repetition of the CS-only showed an astonishing increase in the CR, increasing the blood pressure increment from well below 50 millimetres to something like 250 millimetres. Furthermore Napalkov reports that in some dogs this increase became chronic, illustrating a typical psychosomatic disorder following upon behavioural manipulation. This experiment is illustrative only; many other experiments are reviewed by Eysenck (1976, 1979), some of which were carried out.
before the formulation of the theory, other afterwards, with the specific intention of testing it. On the whole the evidence for the existence of incubation phenomena is strong, although of course the repeated documentation of the existence of the phenomenon does not necessarily endorse the theory formulated to explain it.

Note that the incubation of the CR in the Napalkov experiment follows exactly the path noted for the development of most non-traumatic neurotic disorders, i.e., a relatively slight initial conditioning experience followed by incubation of the CR produced by the presentation of the CS-only a number of times. This is what is usually reported in cases of development of neurotic disorders, and it is strikingly demonstrated in this case in the laboratory study of animal subjects.

The consequences of the presentation of CS-only after Pavlovian B type conditioning are thus rather complex, as this presentation may be followed either by extinction or by incubation. Fig. 4 illustrates how these different consequences can be explained and predicted in terms of the theory. The strength of the CR is given on the ordinate, the duration of CS-only exposure is given on the abscissa. Curve A illustrates that the presentation of the CS-only is at first followed by a very strong CR, but the strength of the CR gradually declines over time until it reaches a very low point. (Note that this curve is identical with that

Fig. 3. Habituation of the UCR, and incubation of the CR after single conditioning trial. Eysenck 1979.
found by Rachman and Hodgson in actual experimental work with obsessive-compulsive patients, as shown in fig. 2). Theoretically we would expect such long term exposure to lead to extinction, so that the next presentation of the CR would follow curve B, i.e. start at a lower point and end at a lower point also. Several evocations would lead to the situation presented in curve C.

The hypothesis presented by Eysenck (1976, 1979) is that there is a critical point on the ordinate, such that above it presentation of the CS only is followed by incubation/enhancement, whereas below it presentation of the CS-only is followed by extinction. The strength of the CR is of course intimately linked with the duration of CS-only exposure; given that the CR is very strong, short exposures will lead to incubation/enhancement, whereas long exposures will lead to extinction. There is much evidence in the experimental and clinical literature to support this hypothesis (Eysenck 1979).

The strength of the CR is of course determined not only by the strength of the UCS, but also by personality factors relating to emotionality/neuroticism, and introversion/extraversion; the general theory linking these personality concepts with neurosis has been explained in Eysenck and Rachman (1965), and application to the present theory in Eysenck (1979). A UCS which is experienced as "strong" by a neurotic or highly emotional person may be experienced as relatively "weak" by a very stable person, and the strength of the conditioned response will depend in part on the conditionability of the subject, i.e. on his position on the introversion—extraversion continuum. The experimentally-based theories relating to these predictions are discussed in Eysenck (1967).

Fig. 4 may illustrate how the general theory here developed unravels certain apparent inconsistencies which exist in the clinical applications of behaviour therapy known as desensitisation and flooding. In the former, the patient is protected against any strong anxiety arising during therapy by a procedure in which he is kept in a relaxed state, and is presented the unreinforced conditioned stimulus, whether in vivo or in imagination, only at points on the hierarchy which are relatively little arousing (i.e. which are well below the critical point, as in curve C in fig. 4). If the critical point is ever exceeded, the CS is immediately withdrawn; it has often been demonstrated that when the critical point is reached or exceeded in desensitisation, the success of treatment is in peril, and the patient is actually made worse (Wolpe 1958). In flooding,
on the other hand, the patient is immediately confronted with the most threatening CS, i.e. the one at the top of the hierarchy; this procedure, which includes explicitly an element of response prevention, is continued for periods of an hour and more, as illustrated in the work of Rachman and Hodgson quoted in an earlier paragraph. Both desensitisation and flooding are successful in practice (Kazdin and Wilson 1978), although they appear to proceed in contradictory directions; this clearly is an anomaly which has posed considerable problems to a theoretical explanation of treatment success.

According to the theory embodied in fig. 4, the answer lies in the short duration of the exposure to high anxiety CS which occurs in desensitisation when the therapist makes an error; the critical point is exceeded and consequently enhancement takes place, rather than extinction, which only occurs at levels of anxiety below the critical point. In flooding, exposure to the CS is continued for a sufficiently lengthy period of time to get well below the critical point; hence extinction takes place and no enhancement.

Fig. 4. Extinction and incubation of anxiety as a function of strength of the CR and duration of CS-only exposure. Eysenck 1979.
Eysenck (1980) has suggested that it is possible to formulate a unified theory of psychotherapy, based on the experimental work and the theoretical developments outlined at the beginning of this paper. According to this theory, all successful treatments of neurosis, including not only the different forms of behaviour therapy (e.g. desensitisation, flooding, modelling, etc.) but also psychoanalysis, psychotherapy, Ellis's rational-emotional type of therapy, Roger's client-centred type of therapy, etc. can be explained in terms of the extinction of conditioned responses; spontaneous remission too, in so far as it is effective, would be covered by this explanation. We have already implicitly or explicitly dealt with the major methods of behaviour therapy, i.e. desensitisation and flooding, modelling also clearly falls within this paradigm. What is done there is to expose the patient to the feared object or situation, keeping his anxiety rather low by having the model cope successfully with the object or situation, and clearly not implicating the patient directly. Much the same comment can be made about Ellis's rational-emotional type of therapy, or Roger's client-centred type of therapy. In both cases what happens is that the patient is exposed, in imagination and through a process of talking, to the feared objects or situations, thus being forced to expose himself to the unreinforced CS, which is in a relatively weak form and thus prevents it from producing CRs above the critical point. The presence of the therapist encouraging the patient and generally serving to relax him, acts as an additional variable reducing the strength of the CR in these situations.

As far as psychotherapy is concerned, it seems clear that something very similar is taking place there also. The patient is encouraged to discuss his difficulties, his anxieties, and to confront the objects or situations producing his fear responses. He does so against a background of a sympathetic, helpful listener who is usually supportive in his attitude, thus reducing generalised anxiety. The critical strength level of the CR is therefore not usually exceeded, and consequently all these encounters should be capable of being symbolised by curve C in our diagram. It is suggested that all successful methods of psychotherapy (and by successful we do not necessarily mean that their effectiveness exceeds that of spontaneous remission, but merely that after therapy the patient is better than he was before therapy) follow this paradigm and are therefore examples of Pavlovian extinction of Pavlovian B type conditioned responses.

Spontaneous remission does not, in our opinion, pose a particular
H.J. Eysenck / Conditioning and neurosis

difficulty to this type of explanation. It is well known that people suffering from neurotic disorders who do not consult a physician, a psychiatrist, a psychoanalyst or a clinical psychologist will consult other people with whom to discuss their troubles — priests, teachers, friends, relatives etc. The conditions for extinction are therefore very similar to those which obtain under psychotherapy, and an exactly similar explanation may therefore be given for the occurrence of extinction. We thus end up with a parsimonious theory which is in good accord with experimental facts ascertained in the laboratory, and which explains all the phenomena of successful treatment — and of successful non-treatment!

It would also seem that our theory can explain the relative success of behavioural therapies and the relatively lower level of success of psychotherapy and spontaneous remission (Kazdin and Wilson 1978; Rachman and Wilson 1980). The methods of behaviour therapy explicitly use the mechanism of extinction, and have been worked out so as to maximise the effectiveness of this mechanism. The methods of psychotherapy have been worked out on different theoretical principles, and these interfere with the quickest method of extinction. The same is probably true of spontaneous remission; the people to whom the neurotic turns have no explicit theory to guide them, and hence their conduct would not be optimal as far as extinction of their conditioned responses is concerned. We would therefore expect, and we do find, that the methods of behaviour therapy tend to work best, followed by psychotherapy, followed by spontaneous remission.

It has been argued (Strupp et al. 1977; Garfield and Bergin 1978) that psychoanalysis and psychotherapy sometimes produce negative effects, i.e. they actually harm the patient. This too can be explained in terms of our theory. Psychoanalysts in particular often adopt a non-helpful, pseudo-objective, interpretative attitude which does not help the patient to relax, or encourage him in any way; under these conditions his anxiety may easily exceed the critical point and thus lead to enhancement rather than to improvement. The evidence presented by Bergin (1963), Bergin and Jasper (1969), Bergin and Solomon (1970), Carkhuff (1967), and Carkhuff and Truax (1965) suggests that the hypothesis linking the personality and the therapeutic manner of the therapist with success and failure along the lines of our theory may be correct, although the evidence they provide is not as firmly established as one would like to think.
It is suggested, then, that, based on experimental evidence from animal and human laboratories, the revised Watsonian theory of conditioning and extinction can explain the development and the successful treatment of neurotic disorders in humans, thus linking together the experimental laboratory and the clinic in a very direct and meaningful fashion. It has always been the contention of the writer than such a connection is essential if psychology is to make a proper scientific contribution to the study of mental abnormality, and it may be suggested that the relative lack of success of psychotherapeutic, psychoanalytic and other mentalistic methods of treatment may be due to their almost complete neglect of the experimental and theoretical knowledge accumulated by academic psychologists in their studies of animal and human subjects.

In recent years, it has become fashionable to discount the successes of the condition/extinction theory in favour of what is sometimes called “cognitive behaviour therapy”. Such a view is based on two quite erroneous premises. In the first place, cognitive features of behaviour are already included in Pavlov’s conception of conditioning, as in his notion of the “second signalling system”. As he pointed out: “A word is as real a conditioned stimulus for man as all the other stimuli in common with animals, but at the same time more all-inclusive than any other stimuli”. And again: “Owing to the entire preceding life of a human adult a word is connected with all the external and internal stimuli coming to the cerebral hemispheres, signals all of them, replaces all of them and can, therefore, evoke all the actions and reactions of the organism which these stimuli produce”. Thus verbal and other cognitive features are not to be considered outside the lawful field of conditioning phenomena.

The other error frequently committed is to imagine that there is in existence some form of systematised set of cognitive rules or laws, commensurate with those known in the conditioning literature. That this is not so has been indicated very clearly by Allport (1975), who concludes his examination of the field by stating that it is characterised by “an uncritical, or 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 mere vacuum of theoretical structure within which to inter-
relate different sets of experimental results, or direct the search for significant new phenomena". Theories of this type are not likely to help us in gaining a scientific understanding of the phenomena of neurosis and treatment.

References

Garcia, J., B. McGovan and K. Green, 1971. 'Sensory quality and integration: constraints on


Les méthodes de la science appliquée sont généralement basées sur les résultats de la science pure, et la découverte de méthodes applicables avec succès dans la thérapie des neuroses dépend de la formulation de théories appropriées pour ces troubles. Il est discuté dans cet article que la seule théorie acceptable pour le moment sur la neurose est une adaptation du modèle de conditionnement de Watson, transformé de telle manière qu’il s’adapte aux découvertes les plus récentes dans les domaines de la formation et de l’extinction de réponses conditionnées. Un tel modèle est d’une grande importance pour la thérapie comportementale et pour les différentes méthodes de traitement des neuroses que recouvre cette dénomination, et évidemment pour expliquer l’efficacité apparente de la psychothérapie, de la psychanalyse, et de tout ce qui peut médier la remission spontanée.