OBJECTIVE PSYCHOLOGICAL TESTS AND THE ASSESSMENT OF DRUG EFFECTS[1]

By H. J. Eysenck

Institute of Psychiatry, University of London, England

I. Introduction 333
II. A Critique of Current Methodology 334
III. A Theory of Depressant and Stimulant Drug Action 342
IV. Experimental Tests in the Assessment of Drug Effects 348
V. A Second Paradigm for the Testing of Psychopharmacological Hypotheses 366
VI. A Mathematical Model of a Psychopharmacological Hypothesis 370
VII. The Place of Animal Work in a Hypothetico-deductive System 375
VIII. A General Dimensional System of Psychopharmacology 379
IX. Summary 382
References 383

I. Introduction

In recent years a good deal of interest has been shown in what has come to be known as psychopharmacology, a term which may be taken to denote the study of the behavioral effects of drugs. While research has been going on in this field ever since Kraepelin tried to put it on a scientific basis, very little of lasting scientific importance has emerged. It is suggested here that this failure to achieve any noteworthy generalizations is due to a lack of theoretical and methodological sophistication, and that what is required in order to improve the situation is a more rigorous theoretical and methodological approach to a subject which, while superficially very simple, is nevertheless bristling with difficulties and complexities not always realized by investigators. The present chapter is devoted to a summary of the theoretical and experimental con-

[1] The work of Micko reported in this paper (and some of the theoretical work also) was supported by the National Institute of Mental Health under grant M-46-59/7590363-93001-01.

siderations and results relating to an attempt to introduce some degree of order into one corner of this chaotic field; no attempt has been made to cover the whole literature on psychopharmacology. Such a summary has been prepared elsewhere (Trouton and Eysenck, 1960), and the interested reader must be referred to this for coverage of the whole field. Interest in this chapter has been rather in certain methodological problems, and experimental studies are quoted more as illustration than for the sake of completeness. I have used one particular hypothesis, and the experimental work relevant to it, to demonstrate how psychopharmacology is related to psychology, how it can derive from it theories and methods for testing such theories, and how without such symbiosis both general psychology and psychopharmacology would be the poorer. Most of the points raised may appear self-evident, but these have seldom been made in print, and in practice they have usually been completely disregarded by experimentalists not brought up in the discipline of psychology. It would appear, therefore, that the contents of this chapter may not be entirely inopportune.

II. A Critique of Current Methodology

Let us consider the usual practice of leading experimenters in this field. We shall take it for granted that in their work such people will not make any of the obvious and elementary mistakes which rule "out-of-court" immediately the majority of papers published (Trouton and Eysenck, 1960). Such errors arise from the possibility of contamination of results through knowledge of the drug administered on the part of either subject or experimenter, i.e., the failure to use the "double-blind" method of administration; the use of unstandardized ratings and other subjective methods of assessment; failure to control and correct dosages and time elements according to the best available knowledge of the drug in question; the use of tests for the measurement of effects which have no theoretical rationale and no evidence of their validity for the particular trait or ability under investigation [the Rorschach test (Eysenck, 1958) may be quoted as an example of this]. Assuming that all the necessary precautions have been observed, we find that in the typical experiment one of two paradigms has been observed,

the choice depending largely on whether or not strong learning effects are present in the test under investigation. According to the first paradigm, which is suitable for tests not showing strong learning effects, half the sample is tested under placebo conditions, the other half under drug conditions; on the second day the first group is now given the drug, the second group is now given a placebo (P). Changes between placebo and drug conditions are evaluated in terms of appropriate statistical techniques. This method and its various modifications will be referred to as method I. Method II, which is appropriate for tests showing strong practice effects, also employs a control and an experimental group, suitably matched on relevant variables; the control group performs the test under placebo conditions, the experimental group under drug conditions. Differences between the groups are evaluated again in terms of the appropriate statistical test which, however, would usually be much less powerful because of the inevitable failure to achieve proper matching between the groups. Thus method II requires considerably greater numbers of subjects than does method I.

Both method I and method II are adequate for the purpose for which they are used, but great difficulties arise in interpreting the results which emerge. Let us assume that a given drug, when compared with (P), has produced significant changes in performance on a group of subjects (S) on test (T). (Cf. Fig. 1 as an illustration.) How can we interpret such a finding? At first sight there appears to be no difficulty. Supposing the test used was the pursuit rotor, i.e., a test in which the subject attempts to keep a metal stylus in touch with a small metal disc set in a Bakelite turntable rotating at a speed of 60 rpm. If a depressant drug is given, performance on this test is lowered. The interpretation usually given would be that: "The drug lowers performance on the pursuit rotor." This, of course, is a true statement of facts, but it would have meaning psychologically only if pursuit rotor performance were determined in its entirety, in everybody, and at all times, by one identical set of abilities and drives. That this is not so is shown very clearly in the experimental work of Fleishman and Hempel (1954, 1955), Fleishman (1956). These authors showed, in brief, that when scores made by different subjects on the pursuit rotor at different points on the learning curve (i.e., after different periods of practice) are intercorrelated with performance scores on a large

variety of other motor tests, then it is found (*a*) that performance at any one point is a compound of different abilities, i.e., performance on the pursuit rotor has significant loadings on several different motor ability factors; and (*b*) that the spectrum of abilities

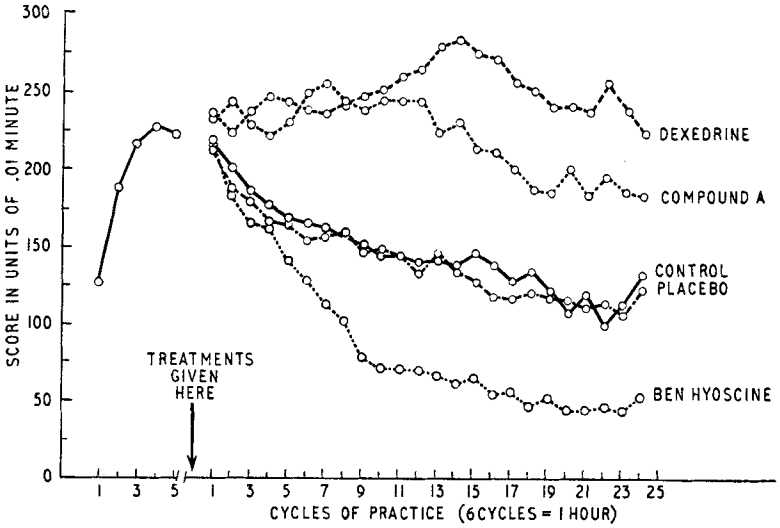

FIG. 1. Effects of administration of stimulant and depressant drugs upon massed practice. Taken from Payne and Hauty (1955).

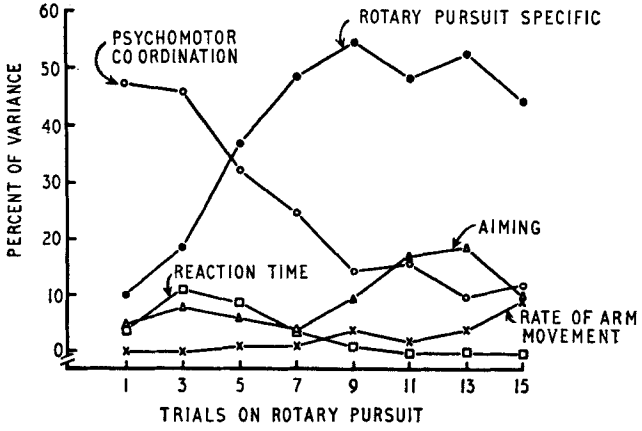

FIG. 2. Diagram illustrating the different factorial composition of performance on the pursuit rotor at different stages of practice. Taken from Fleishman (1956).

which goes to make up the factor loadings of pursuit rotor performance changes in the course of practice, i.e., pursuit rotor performance may have a high loading on factor I and a low loading on factor II at the beginning of practice, and a low loading on

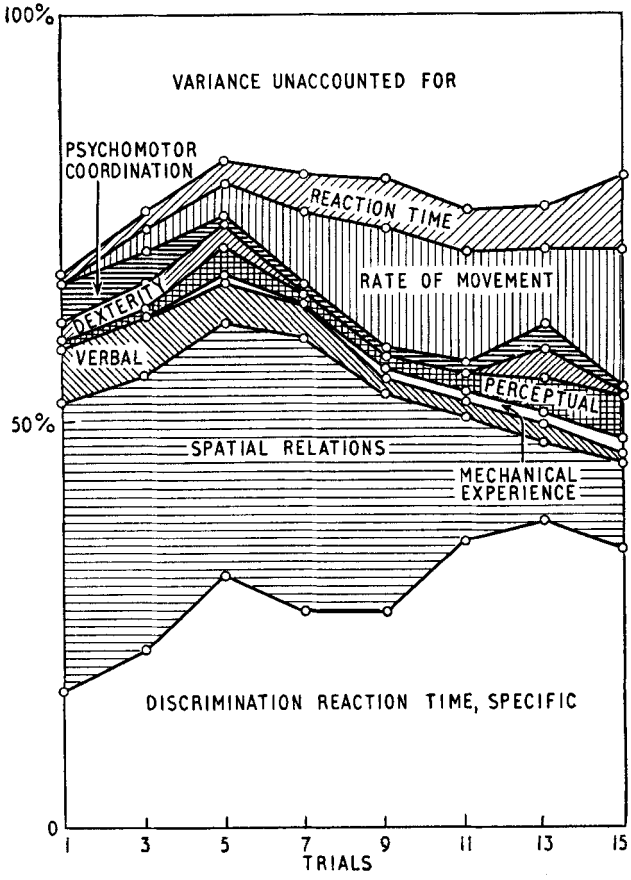

FIG. 3. Diagram illustrating the different factorial composition on discrimination reaction time task. Taken from Fleishman (1956).

factor I and a high loading on factor II at the end of practice. Figure 2 shows in diagrammatic form some of Fleishman's findings. Figure 3 shows the same phenomenon with respect to another widely used test in psychopharmacological work, namely reaction time experiments.

We may put the matter in a mathematical way by writing:

$$\sigma_P^2 = \sigma_{F_1}^2 + \sigma_{F_2}^2 + \ldots \sigma_{F_n}^2 + \sigma_S^2 + \sigma_E^2$$

in which σ_P^2 stands for the total variance on the pursuit rotor test, while F_1, F_2 . . . F_n stands for different abilities; P is performance, S stands for the specific factors involved in the test, and E refers to the error variance. (The total equation refers to pursuit rotor performance at any one moment of time only.) Now consider the effect of drug on performance. The effect of the drug could be on any one element or any combination of elements in the equation; the simple fact that total performance is lowered tells us nothing whatsoever about the effect of the drug on those elements which are of real concern to us, i.e., the different abilities F_1, F_2 . . . F_n in the formula. The result of any such experiment, therefore, is completely equivocal, and cannot be directly interpreted.

The situation is worse, of course, when we consider that the equation given above only holds for one particular level of training, and changes as practice proceeds. Thus, even if we knew the effect of the drug on the different constituents of the equation at one level of training we would not necessarily be able to say anything about other levels of training. A particular example may make this clear. At the beginning of practice on the pursuit rotor the subject spends a lot of time hunting for the disc with his stylus (psychomotor coordination) and hardly any time in simple rotary pursuit movement of the arm. Let us arbitrarily call the abilities involved in these two activities F_1 and F_2. When the subject has had a good deal of practice he spends very little time on activities of the kind F_1 and more time on activities of the type F_2. Now if the drug happens to affect F_1 but not F_2, then we might find systematic changes in pursuit rotor performance in the sense that at the beginning, where F_1 played an important part, performance was impeded, whereas later on, when F_2 took over, performance was not impeded by the intake of the drug. Thus two different investigators, starting subjects at different levels of practice, or using discs of different size (thus producing different levels of difficulty), might easily obtain contradictory results, and indeed the frequency of contradictory results in this field suggests that the above considerations may be of some importance. It is quite possible that a given drug may have a deleterious effect on F_1; may improve F_2,

and may leave factor F_3 unaffected; under those conditions any experimental result is possible depending on the degree of practice the subjects have had. (Further complications are introduced, of course, by the fact that one ability may often be substituted for another in achieving a given performance score, thus making the analysis even more complex. There is not room in this brief review to deal with complexities of this kind.)

The problem suggested by these considerations can be solved by studying the effects of drugs not on an isolated test performance, but rather on factors derived statistically from whole patterns of tests. There are obvious practical difficulties in the way of this solution, and it would certainly be far more expensive, time-consuming, and complicated than the much simpler type of experimental procedures currently employed. Nevertheless, the writer is firmly convinced that the gain in rigor, and in interpretability and duplicability of results, will eventually enforce some such solution, and will relegate the simple "one test experiment" to the field of preliminary investigation.

The difficulty discussed above arises from the complexity of any particular test that may be chosen to measure drug effects. Other difficulties in interpretation arise from the fact that performance on a given test, even if this should be completely unifactorial, is, nonetheless, determined by a number of different parameters which must be specified. Such specification has usually been the task of learning theory. Let us consider again performance on the pursuit rotor, and let us assume that performance of the drug group falls off as compared with the performance of the placebo group. Let us also assume that we are dealing merely with one ability, say F_1. Now in terms of learning theory we find that the following elements should be taken into account in assessing the causes of this decrement. (1) There may have been a decrement in what is known in learning theory as ${}_sH_R$ or excitatory potential, i.e., the ability of the central nervous system (CNS) to form the new connections which are the neurological basis of learning. (2) There may have been an increment in what is known in learning theory as I_R or reactive inhibition, i.e., neural fatigue affecting those cortical elements underlying the passage of neural impulses during the performance of the task; reactive inhibition is not assumed to be permanent like ${}_sH_R$, but to

dissipate in time. (3) Alternatively there may have been a decrement in drive symbolized in learning theory as D; this is supposed to interact with habit in multiplicative fashion to produce ${}_{S}E_{R}$ or performance. (4) Lastly, there may have been an increment in conditioned inhibition or ${}_{S}I_{R}$ which arises after a long period of massed practice. (As this rather complex variable need play no part in this example or in the remainder of this chapter, I will not try to explain it in detail here. A brief account of the relevant aspects of modern learning theory is given in Eysenck, 1957b.)

Expanding the general formula linking these various concepts, we get ${}_{S}E_{R} = (D \times {}_{S}H_{R}) - (I_{R} \times {}_{S}H_{R})$, so that if we have no reason to suspect D of being affected either we must conclude that the drug has decreased ${}_{S}H_{R}$ or that it has decreased I_{R}, or both. Clearly the experiment which leads to the simple report that a drug has decreased performance on pursuit rotor learning, even under the simplified conditions assumed to hold here, is by no means unequivocal; instead of answering a psychological problem, it raises one, and leaves us to wonder which of the several possibilities mentioned above is the correct one. Fortunately the theory in question also enables us to design an experiment to answer this particular question. As mentioned above, I_{R} dissipates with time, whereas ${}_{S}H_{R}$ does not. This gives rise to so-called reminiscence effect; after a period of massed practice, during which I_{R} has been built up, a rest pause will serve to dissipate this I_{R}; consequently after the rest pause performance is distinctly superior (because of the elimination of I_{R} which acts to depress performance) as compared with prerest performance. This sudden improvement after rest is referred to as the reminiscence effect. If now the drug acts on I_{R}, then a comparison may be made of two groups, both tested on 2 successive days. The experimental group would be given the drug on the first day and (P) on the second day, the control group would be given (P) on both days. If the drug affects I_{R}, then a significantly greater reminiscence effect should be found for the experimental group. If, on the other hand, it is ${}_{S}H_{R}$ which is affected, no such reminiscence effects should appear and performance of the experimental group on the second day should remain inferior to that of the control group. As an illustration we may quote the data in Fig. 4 which were collected in an unpublished experiment by Micko. Using the pursuit rotor,

he tested twenty-five subjects each under 250 mg Doriden (D) and twenty-seven subjects under (P) conditions during a 5-minute practice period on the first day. Both groups were retested on the second day, again for 5 minutes; this time both groups were under no-drug condition. Scoring was carried out for successive 10-second periods, and the time on target during this 10-second period is plotted on the ordinate. It will be seen that during the first day the (D) group is inferior to the (P) group for almost the whole of the 5-minute period. Because of the great individual differences, however, this difference does not achieve significance. When we

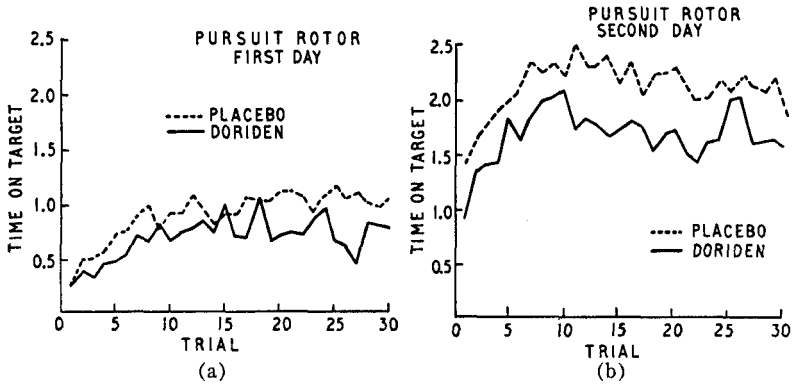

FIG. 4. Performance on pursuit rotor under placebo conditions and after administration of Doriden.

turn to the second day, we find that now the two groups are even more clearly distinguished and the performance of the (D) group is found to be significantly lower at the 1% level on the first trial. As this is a crucial trial for reminiscence, it can be seen that the results favor the hypothesis that (D) acts on ${}_{S}H_{R}$, reducing the excitatory potential to a significant degree.

A similar study is illustrated in Fig. 5 in which are plotted the results from an investigation of eyeblink conditioning, carried out on the same subjects as the experiment just mentioned. It will be seen that during the first day the (P) group is superior to the (D) group, the mean number of responses being, respectively, 3.32 and 4.15, giving a difference of 0.83. Differences between the two groups on the second day are illustrated in full, and the degree of superiority of the (P) group is indicated by the size

of the cross-hatched area. On the second day the mean number of responses was 4.96 and 5.89, respectively, giving a difference of 0.93. Again, any individual differences are so large as to make the effect of the drug of doubtful statistical significance, but the inadequate data suggest again that it is ${}_{S}H_{R}$ or excitatory potential which is most affected by (D).

These data are illustrative and are not meant to suggest any definitive conclusions. They are merely meant to show that under the appropriate experimental conditions the problems raised in this section can receive an answer. But again it is essential that we get away from the simple type of paradigm outlined at the

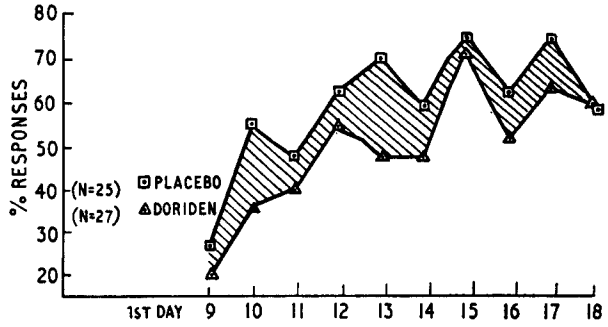

FIG. 5. Proportion of responses on eyeblink conditioning apparatus under no-drug condition of subjects who on the previous day had been given eight conditioning trials under either placebo or Doriden administration.

beginning of this chapter and resort to a more sophisticated type of experiment, the details of which are dictated by the most advanced and experimentally supported type of theory available in psychology.

III. A Theory of Depressant and Stimulant Drug Action

We have seen in the preceding section that measures of the behavioral effects of drug actions, in order to be interpretable, should be part of a general psychological theory. It is the main purpose of this chapter to show how such a theory might run, and how support for it may be found in experimental studies which shall not fall foul of the rules suggested in the first section. The

theory to be discussed is one which has links both with learning theory and with personality theory; these links have been discussed in great detail elsewhere (Eysenck, 1957a,b) and will only briefly be referred to here.

We have mentioned in the preceding section the two notions of excitation (leading to ${}_{S}H_{R}$) and inhibition (leading to I_{R}). Pavlov originally suggested that different organisms differ with respect to the strength of these two variables in such a way that individuals could be arranged along a continuum depending on their excitation-inhibition balance; at the one extreme would be subjects strong in excitation and weak in inhibition, at the other extreme those strong in inhibition and weak in excitation. More balanced types would be intermediate. This notion was taken up by the writer (Eysenck, 1957a,b) as a possible basis for the personality dimension of extraversion-introversion, in the sense that the person strong in excitation and weak in inhibition would thereby be predisposed to introverted behavior patterns, while the person strong in inhibition and weak in excitation would be predisposed to extraverted behavior patterns.[2] It was also proposed that this general rule would hold not only for normals but also for neurotics; it had been shown previously that hysterics and psychopaths tend to lie toward the extraverted end of the continuum, whereas dysthymics (anxiety states, reactive depressives, obsessionals, etc.) tend to lie toward the introverted end of the continuum. Furthermore, evidence was given to show that brain damage, particularly the operation of leukotomy, has an extraverting effect on personality, and may be assumed to increase inhibitory potential and to decrease excitatory

[2] A detailed discussion of the concepts of extraversion and introversion has been given by the author elsewhere (Eysenck, 1959d). There are many different ways of diagnosing the position of a particular person on the extraversion-introversion continuum. It may be done in terms of ratings, questionnaire responses, psychiatric diagnoses (psychopaths and hysterics tend to be extraverted neurotics; anxiety states, obsessionals, and reactive depressions tend to be introverted neurotics); objective tests of the "miniature situation" kind; physiological tests; and objective performance tests, linked in a rational manner to the general theory of extraversion-introversion (Eysenck, 1957b). It should be noted that these concepts are operationally defined in terms of the instruments used for measuring them and that, while descriptively they are similar in many ways to Jung's terminology, there is no intention of accepting Jung's psychoanalytical and somewhat mystical theory regarding the nature and origin of these personality types.

potential. We thus have a complex of relationships linking the excitation-inhibition balance to personality, both normal and abnormal, and to brain damage.

For proof of hypotheses such as those just suggested, we require tests which can be regarded as measures of extraversion-introversion on the one hand, and measures of the excitation-inhibition balance on the other. As extraversion and introversion are behavior patterns observable through life, we can obtain estimates of an individual's position on this continuum either in terms of ratings, psychiatric or otherwise (Eysenck, 1959d), or by means of self-ratings on specially constructed scales, such as the Maudsley Personality Inventory (MPI) extraversion scale (Eysenck, 1959a). As regards the excitation-inhibition balance, we must look for measures of this in the field of general and experimental psychology, particularly in relation to experiments for the explanation of which the concepts of excitation and inhibition have been used. Pavlov himself used salivary conditioning in dogs as a measure of this balance; it is clear that dogs strong in excitation and weak in inhibition should condition well and quickly, and should retain their conditioned responses for a long time, while dogs strong in inhibition and weak in excitation should condition only weakly and with difficulty, and should retain their conditioned responses only for a short while. Many other measures relevant to the excitation-inhibition balance have been proposed (Eysenck, 1957b), and some of these will be discussed in more detail when we turn to a consideration of the effects of drugs on these measures.

We are now ready to define the relationship between groups of drugs and variables considered above. Put briefly, the writer's drug postulate states that: "Depressant drugs increase cortical inhibition, decrease cortical excitation and thereby produce extraverted behavior patterns. Stimulant drugs decrease cortical inhibition, increase cortical excitation and thereby produce introverted behavior patterns." We have already discussed the meaning of excitation, inhibition, extraversion, and introversion in this postulate; we must now discuss the meaning of "depressant" and "stimulant." These terms are often used by pharmacologists to group together drugs such as alcohol, hypnotic and tranquillizing drugs, barbiturates, and so forth, on the one hand, and caffeine and the amphetamines, generally on the other. This pharmacological group-

ing is accepted as a first rough guide to the selection of drugs which would form the two groups included in the drug postulate, but it should be noted very clearly that ultimately the terms "depressant" and "stimulant" are used in a psychological context and must derive their definition from the behavioral effects of the drugs concerned. We might therefore say that in the long run the postulate may be used as a definition of depressant and stimulant drugs, respectively, in the sense that all drugs which increase cortical inhibition and decrease cortical excitation form the group of depressant drugs, while conversely, all those drugs which decrease cortical inhibition and increase cortical excitation form the group of stimulant drugs. The present pharmacological grouping is merely a suggestive first step toward a proper psychological grouping in terms of this drug postulate.

It might be objected that this is a circular process, in which drugs are defined as belonging to two great groups in terms of their effects on behavior, and that then the behavior effect of these drugs confirms the quoted hypothesis. Such an objection would be true but irrelevant; our interest lies in the possibility of experimentally defining an invariance of the kind outlined in the postulate; the discovery of such an invariance is of considerable interest and importance, although like most scientific advances it does involve a circular argument.

If the invariance suggested should in fact be experimentally supported in a number of cases, then we might even be able to invert the argument and say that if a drug known to be a depressant drug according to our psychological definition has a certain effect on a given test, then this test is *eo ipso* determined in part by cortical excitation or inhibition and may therefore be regarded as a measure of extraversion-introversion. It would, of course, be necessary to consider very carefully and critically side effects and other possibilities of erroneous inference, but by and large the argument from the invariance obviously can be taken in both directions. There clearly is a vast amount of experimental work waiting to be done before such a postulate can be regarded as firmly established and before we can use it effectively.

Before turning to a consideration of the evidence, one example may be given in some detail to suggest the correct method of using the drug postulate. There is a well-known effect which

emerges in studies using serial learning of nonsense syllables, i.e., studies in which series of nonsense syllables are shown to the subject for 2 seconds at a time, to be learned by him in such a way that he can anticipate the next syllable while the previous one is being shown. Series may be of any length, but usually consist of 8, 12, or 15 syllables, and a variable amount of time, usually 6 seconds, elapses between series. It is usually found that when the position of the syllables in the series is plotted on the abscissa and the number of repetitions necessary before a syllable is learned is plotted on the ordinate, then the resulting curve is bowed upward in the center, suggesting that the central syllables take longer to learn than those at the beginning or the end of the series. This effect is usually explained in terms of the Hull-Lepley hypothesis (Hull *et al.*, 1940) as being due to trace inhibition developing between syllables within the series, and being most extensive at the center, thus producing the bowing there. If this hypothesis were true, then it would follow from our postulate (*a*) that extraverts would show greater bowing effects than introverts, and (*b*) that depressant drugs would increase and stimulant drugs decrease the bowing effect. Figure 6 shows the results of one test of this hypothesis, in which twelve nonsense syllables were used and four groups of subjects tested under (P), Amytal (A), and Dexedrine (D-A) conditions; testing after (D-A) was begun either relatively soon or relatively late after administration (hence the groups were called "short (D-A)" and "long (D-A)" (Eysenck, 1957b, quoting unpublished work by Dr. Willett). It will be seen that there are no differences in the bowing effect between the groups, and at the time when the results were reported the writer pointed out that: "This failure may mean one of three things: (1) The drug postulate is in error. (2) The Lepley theory, according to which the bowing of the serial position curve is due to inhibition, is in error. (3) Drug effects are too slight to give demonstrable results in this situation." As other tests had shown quite large drug effects, the third possibility is not a likely one, and it seems desirable to test the second possibility by designing a direct test of the Hull-Lepley theory. This was done by getting subjects to learn a series of nonsense syllables without interposing any breaks between series; in this way the pattern of trace inhibitions spanning the within-series syllables only would be broken and the bowing

effect should therefore be largely destroyed. In actual fact no such destruction was observed and the bowing effect was as strong as ever, thus showing effectively that the Lepley hypothesis was wrong and the bowing effect not due to inhibition as postulated (Eysenck, 1959e). It follows that the test is not a proper test of the theory and does not invalidate the drug postulate. It also follows that we would not now expect extraverts to have greater bowing

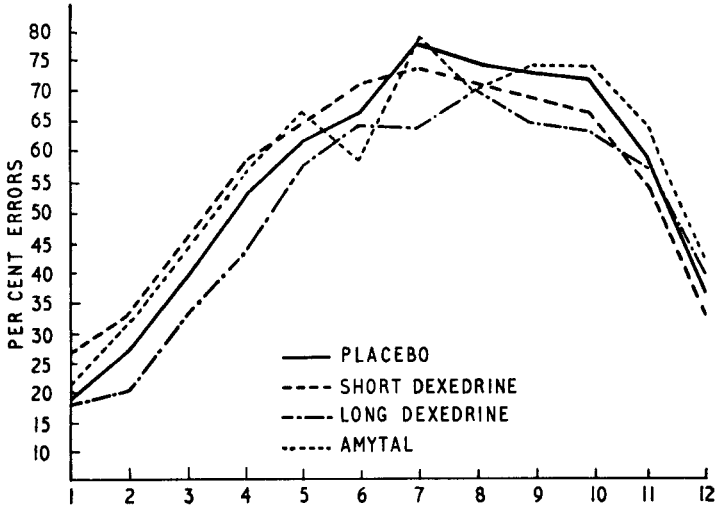

FIG. 6. Nonsense syllable learning: serial position curves of groups of subjects administered different drugs.

effects than introverts, and indeed no differences between extraverts and introverts have been found on this measure. This example shows the kind of reasoning applied in this situation, and it also shows the necessity of being careful in interpreting results, particularly where these are negative. It is always possible that a negative result is due not to a fault in the postulate but rather to an erroneous deduction due in turn to a fault in the general psychological theory of the phenomenon being used as a test. Great care must obviously be taken not to over-interpret data at such an early stage of development of psychological theory (Eysenck, 1960).

IV. Experimental Tests in the Assessment of Drug Effects

In this section we will deal systematically with experiments carried out in the field of learning and conditioning, motor performance, perceptual performance, and autonomic reactions. This is not a survey of all the available evidence; such a survey has been given by Trouton and Eysenck (1960). We have concentrated very much on work done in the writer's department with the specific intention of testing the hypothesis outlined above; work of others has only been drawn in occasionally when it seemed particularly relevant. All the work quoted relates to experiments performed in the laboratories; studies dealing with observation of behavior (Laverty and Franks, 1956) have not been dealt with here even though they were carried out in an attempt to test the theory in question.

1. *Eyeblink Conditioning*

. In view of the fact that Pavlov's original theory was concerned largely with conditioning it seems appropriate to begin our work with a conditioning measure, more particularly as this had also been used to verify the hypothesis linking extraversion and introversion to the excitation-inhibition balance. The first study was carried out by Franks and Laverty (1955) on sixteen neurotic hospitalized subjects. Four different treatments were used, two of these consisting of different doses of (A), one of (P), and one of no drug or (P). The subjects were given various personality scales including the Guilford R scale, which has been found to be a good measure of extraversion. The experimental test used was one of eyeblink conditioning, carried out in a soundproof conditioning laboratory. The unconditioned stimulus was a puff of air, delivered to one eye at a pressure of appoximately 65 mm of mercury from a distance of 2 cm from the eye and lasting 500 msec. The conditioned stimulus was a pure tone of frequency 1,100 cycles per second at an intensity of 65 decibel above the subject's auditory threshold. The duration of this tone, heard through a pair of padded earphones was 800 msec. The air puff was so arranged that it began 350 msec after the tone had commenced. Partial conditioning was used, the sequence being such that the reinforce-

ment ratio was approximately 60% throughout the reinforcement trials. Thirty reinforcement trials were given, interspersed with eighteen test trials, consisting of the conditioned stimulus alone. The intertrial interval varied from 20–30 seconds with a mean of approximately 25 seconds. After the thirty reinforcement trials and eighteen test trials had been given the subjects were given a further series of ten consecutive test trials (called extinction test trials) so that the resistance to extinction could be measured.

The results showed the following number of conditioned responses during acquisition:

(A), dose of 300 mg = 2.0; (A), dose of 90 mg = 4.0; (P) = 7.2; (N), control run = 7.9.

The number of conditioned responses during extinction for the four groups in the same order was 0.1, 1.9, 3.8, and 3.9, respectively. Extraversion scores under the four treatment conditions were 33.5, 37.8, 22.7, and 23.2. The authors conclude that "The two main hypotheses of this pilot study would seem to be confirmed tentatively. These are (1) that intravenous (A) reduces the number of conditioned eyeblink responses during acquisition and increases the rate of extinction; (2) that intravenous (A) increases the extraversion score as measured by Guilford's R scale, . . . Under the conditions of the present study a dosage difference of 2 grains intravenously produces no significant differences in the measures under investigation. The trends in the data however are consistent with the possibility that eyelid conditionability is some undetermined negative function of the size of dose administered."

In a later study, Franks and Trouton (1958) repeated the experiment using as subjects eighty paid volunteer graduate female students of education, and employing (A) and (D-A) as well as (P), as the drugs to be investigated. The results of the experiment are shown in Fig. 7. The (D-A) group shown in this figure had been administered the drug 2 hours beforehand; another group which had been administered the drug 45 minutes beforehand showed no difference from the (P) group. The authors conclude: "It was found that the 2-hour (D-A) group conditioned more readily than the (P) group, whereas the (A) group conditioned less readily."

The results of the two experiments are in good agreement

with each other as well as with those of earlier writers from Pavlov onward. They are also in good agreement with the drug postulate in showing that the depressant drug does in fact have an extraverting effect while the stimulant drug has an introverting effect. The data leave unanswered the question of whether the effects are produced by acting on the excitatory or the inhibitory potential, or both; this point has already been discussed in a previous section.

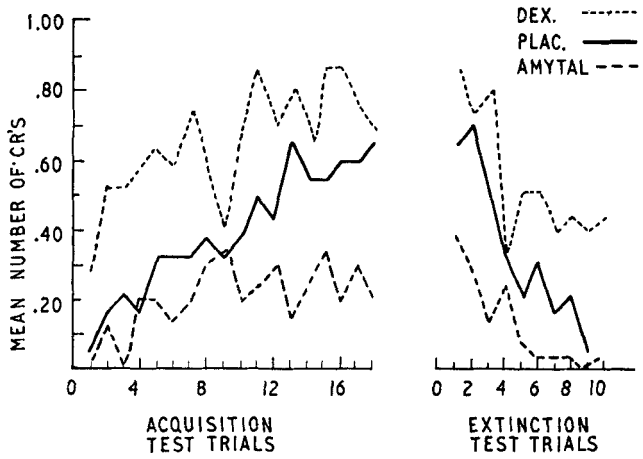

FIG. 7. Acquisition and extinction of eyeblink conditioning in groups of subjects administered different drugs. From Franks and Trouton (1958).

2. *Nonsense Syllable Learning*

Several experiments have been carried out in the field of nonsense syllable learning. Some of these were related to the bowing of the serial learning curve, and this point has already been discussed. Two other predictions were tested by Willett (Eysenck, 1957b), namely: (1) that the mean number of anticipatory errors is increased by (D-A) and decreased by (A), and (2) that the number of repetitions needed to reach the criterion of learning is increased by (A) and decreased by (D-A). Six subjects in all were used, and tested under (P), (A), and (D-A) conditions; testing under (D-A) was carried out twice, once with short, the other time with a long interval between the administration of the drug and the carrying out of the test.

The mean number of anticipatory errors as a percentage of the

mean trial number was, as predicted, lowest for (A) (38.99%) intermediate for (P) (41.31%) and longest for (D-A) (48.13%). These figures refer to the condition under which a long time elapsed between the administration and testing; when (D-A) was administered shortly before testing, the figures were quite out of line with prediction (36.08%). As in our other experiments we have also found this condition to give results less in accordance with prediction, the possibility cannot be ruled out that a lengthy period has to elapse before (D-A) becomes fully effective.

As regards the number of trials to learn the list of twelve nonsense syllables to the criterion, the following results were obtained. Under (A) number of trials was 40.68; under (P) conditions it was 37.21; and under (D-A), short and long, respectively, it was 34.00 and 34.07. The results, while in line with prediction, fall short of significance when tested by analysis of variance.

More recently Willett (1960) carried out another similar experiment using twenty-four subjects who were tested under (D), meprobamate (M), and no-drug conditions. The main purpose of the experiment was to test the rate of learning which it was predicted would be depressed by depressant drugs. The number of trials taken to achieve the criterion just failed to achieve significance, but when the scores of individuals were plotted, it became apparent that the resulting slopes were in the form of a simple exponential function. The score on each trial was then put into log form so that the exponential groups could be plotted as straight lines whose characteristics would be readily calculable. These derived slope values were submitted to an analysis of variance, and it was found that the mean slope of the no-drug group was significantly steeper than that of either of those of the two drug groups. Since the slopes of learning curves represent rates of learning, this result means that both (D) and (M) have decreased the rate of learning of the nonsense syllables. On the whole then the results of these studies appear to support the general hypothesis that nonsense syllable learning is affected in predictable ways by depressant and stimulant drugs.

There is one study by Burstein and Dorfmann (1959) which appears to contradict this statement. They examined the effect of (M) on the learning of a list of paired associates of high interitem competition. They paired the performance of some sixty subjects

divided into a (P) group and a group receiving 1200 mg of (M), and found at a 5% level of significance that the (M) group learned the list more rapidly. They explained this in terms of its action in reducing the anxiety level since, according to the well-known Spence-Taylor hypothesis (Eysenck, 1957b; Taylor, 1951), high anxiety will interfere with the learning of complex material. However, Summerfield and Steinberg (1957) have shown in a well-executed study that in comparing the retroactive interfering effect of an interpolated task accompanied by, and without the administration of, a depressant drug, this drug reduced the commonly observed decrement when performance of the original task was resumed. This result they put down to the drug inhibiting the process of interference during the interpolated task. It might equally be argued, therefore, in the case of Burstein and Dorfmann, that the action of the (M) in their study was rather one of inhibiting interference in a manner similar to that described by Summerfield and Steinberg. A direct test of this explanation would seem to be desirable.

3. *Pursuit Rotor Learning*

In the experiment now to be reported the pursuit rotor was used in the form described in a previous section. Four groups of subjects took part in this experiment; these were identical with those tested by Franks and Trouton in the experiment on conditioning described above and were given the same four treatments (Eysenck *et al.*, 1957a). They performed for a 5-minute period consecutive practice on the pursuit rotor; this was followed by a 10-minute rest, a second 5-minute period of continuous practice, a second 10-minute rest, and a third 5-minute practice period. It was predicted that the (A) group, due to either an increase in inhibition or a decrease in excitation, would perform worse and worse as time went on, while the (D-A) group for the opposite reasons would be continuously more superior to the (P) group, which would be expected to be intermediate between the others. Results are shown in Fig. 8, which on the whole bears out these predictions. The "long (D-A)" group is distinct from the others even in the first 5 minutes of practice. During the second 5 minutes of practice the (A) group also is separated out. It is not, however, until the short period of practice that the "short (D-A)" group is

separated from the (P) group. An analysis of variance showed significance for both periods and treatments as well as for the period-treatment interaction. "We may conclude therefore that the drugs used have differentiating effects on performance, that these effects are in line with prediction and that they interact

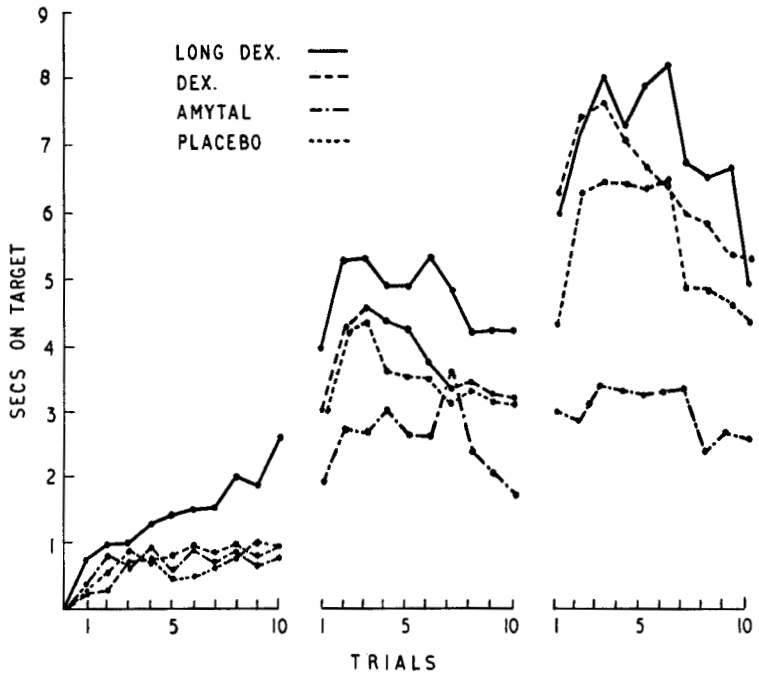

FIG. 8. Performance on pursuit rotor of groups of subjects administered different drugs. From Eysenck *et al.* (1957a).

in complex ways with stage of practice and/or time elapsed since administration."

It is interesting to note that not only do the drugs have the predicted effect, but also that it can be shown that extraverts tend to be inferior to introverts as practice goes on. Evidence for this has been shown in the paper by Eysenck and his co-workers (1957a). It may also be of interest to note that the findings of the drug study are in good agreement with the similar work by Hauty and Payne (1955) and Payne and Hauty (1953, 1955).

4. *Reaction Times*

The action of drugs on reaction times has often been studied (Trouton and Eysenck, 1960), but from the theoretical point of view it is by no means clear what kind of predictions one would make from the theory elaborated in this chapter. There seems no obvious reason why depressant and stimulant drugs per se should affect the simple reaction time with a regular warning signal, provided that the warning signal precedes the stimulus only by a very short amount of time. On the other hand, in experiments without a warning signal, or in experiments where a warning signal is either irregular or precedes a stimulus by a long period of time, we might predict that the impaired vigilance of the extravert or the person subjected to depressant drug treatment would produce longer reaction times and a greater variability of reaction time. The argument is more fully stated in the section on vigilance; the prediction made here will be seen to apply to reaction times only in so far as the simple straightforward reaction time experiment is complicated to take on the character of a vigilance test.

An experiment to test this hypothesis was performed by I. Martin (1960) using twenty-four subjects who were given (M) and (D), and were also tested under a control condition. Two series of tests were done; in the regular series the reaction times were obtained to twenty tones which were regularly preceded by a warning tone at an interval of 3 seconds. In the irregular series the twenty signal-to-respond tones were preceded by a warning tone at a time interval which varied from 3 seconds to 18 seconds. In both conditions a period of about 15 seconds elapsed between the subject's response and the presentation of the next preparatory tone.

Results were evaluated by means of an analysis of variance and it was found that the only significant effect of the drugs occurred during the irregular presentation, and only under the effect of 250 mg (D). There was a lengthening of reaction times during the regular series under (D) but this was not significant. It must be concluded that the hypothesis under investigation is partly confirmed by these data, although it is not clear why 400 mg (M) had no apparent effect on the reaction times under either condition.

5. *Kinesthetic Figural After-Effects*

Prolonged bombardment of sensory receptors sets up a form of inhibition which is known as satiation, and which can be measured by means of what Kohler has called figural after-effects (McEwen, 1958). These may be visual, auditory, or kinesthetic, and they consist in the displacement of contours, or in the changing of sensations following a lengthy inspection period of some object presented to the appropriate sense. It follows directly from the general hypothesis here investigated that satiation should be stronger in extraverts, or after depressant drug administration, while it should be weaker in introverts or after stimulant drug administration. The only study published to date on the effects of drugs is one by Poser (1958) in which he used a measure previously employed by the writer to show that the predicted personality differences between extraverts and introverts did in fact occur (Eysenck, 1957b). Poser tested three groups of ten male volunteers before and 2 hours after administration of (A), (D-A), and (P), respectively. "The prediction that (D-A) would reduce satiation effects was borne out very significantly. The effect of (A) was however indistinguishable from (P) effects." It appears, therefore, that the predicted effects tend to occur, although the failure of (A) to have a significant effect requires explanation.

6. *Apparent Movement*

It is well known that if two stimuli are exposed in rapid succession and in close propinquity, then apparent movement may be observed to occur from the one to the other; the apparent movement seen when stationary pictures are projected in the form of film are well-known examples of this. There is no widely accepted hypothesis relating to the explanation of this phenomenon, but it is usually assumed that a certain amount of interaction and irradiation occurs between the two foci in the cortex, which may be presumed to represent the respective receiving centers of the two original images. If we may further assume that the degree of inhibition of the intervening cortical structures determines the speed of transmission from one of these foci to the other, then we can predict that the threshold for apparent movement will be lowered by depressant drugs, in extraverts and in brain damaged

subjects. This prediction was tested by Costello (1960a) in an experiment in which six subjects were tested twice, once under (P) and once under 600 mg of (M), testing on each day being done once before administration of the drug and 1, 2½, 3½, and 4½ hours after administration. The Wither's direct exposure tachistoscope and an electronic cycling timer were used. The exposure time was 45 msec, the stimulus figure used was a simple one, consisting of two circles of light 1½ cm in diameter and 3¼ cm apart. After a number of practice trials the subject was instructed to report whenever he saw a change from simultaneity to movement or from movement to simultaneity (this threshold was used because there were difficulties in obtaining a satisfactory lower threshold, i.e. one between succession and movement). The time interval at which the subject reported change was reported in milliseconds. Four trials were given, with alternating ascending and descending trials. The means of the four trials constituted the threshold for that session.

The results were treated by analysis of variance and it was found that two significant *F* ratios emerged, namely those for subjects and treatments. As predicted, (M) lowered the threshold as compared with (P) conditions, i.e. the change from apparent simultaneity to apparent movement occurred at longer time intervals under (M). Time of testing apparently had no effect and neither was there any significant subject-treatment interaction. The results on the whole support the hypothesis under investigation.

7. *Visual-Field Effects*

In spite of the constant use of perimeters and other devices for measuring the visual field, there has been very little study of the effect of drugs in this connection. It has been found that anoxia acts to reduce the extent of the visual field, but this may be due to a generally depressant effect of anoxia on responses. It would appear that the hypothesis with which we are dealing here would be capable of mediating certain predictions because it implies that the excitation and extinction processes which characterize the specific isopters will be changed in such a way as to alter the extent and gradient (slope) of visual sensitivity. An inhibitory drug should reduce and depress the field, whereas a stimulant drug should enlarge and raise it. A specific study to test this prediction was

carried out by H. Holland (1960) using a Goldman spherical projection perimeter, in which bowl and target lighting are derived from the same source. The bowl lighting is variable over a very wide range and the bowl-target ratio is achieved by filter and photometer screen. In the present investigation the target size was kept constant at 0.25 mm^2 and the light at 358 lux which gave, with a bowl albedo of 0.7, a reflected intensity of 250 apostilbs.

Only the left temporal field was investigated. Inner and outer thresholds were determined along 14 meridians, every 15° with the exception of the horizontal coordinate which was missed, and readings taken at approximately 83° and 92°, respectively. Three conditions of testing were employed, namely, no drug, (M), and (D). The two latter were used as depressant drugs. Twenty-four subjects in all were employed in the investigation.

Results differed very much for incoming and outgoing movement of the target. The action of the two drugs appears to reduce the excitatory value of the incoming target but to increase it for an outgoing target. If the field is defined in terms of the inner limen, results determine a clear-cut reduction in sensitivity, as predicted. On the other hand, if they are defined by the outer limen they indicate an increase in sensitivity. Finally, if they are defined by the mean of the two, the target isopter, they indicate no significant change at all. These results are highly significant by analysis of variance, and in particular the difference between inner and outer limen produces very marked differences between (P) and drug conditions. Holland points out that "As the inhibitory hypothesis advanced originally would not account for these findings, for it implies that excitation and extinction will co-vary, any explanation must be *post hoc* and speculative. At least two, however, can reasonably be advanced. The first such explanation suggests that there is some change in the psycho-physical scale, whereas the second postulates a general slowing of response, which, within the framework of the experimental procedure, gives a false picture of changed sensitivity. It would be tempting to accept the second type of explanation, which involves a notion of some kind of mental or perceptual inertia, and to conclude by suggesting methods whereby experiments of a similar nature could control for such a possibility. However, to do so would ignore the fact that several other tests involving rapid responses have been administered to the same sub-

jects and have failed to lend support to it. . . . Apart from these facts, the concept of mental inertia is a poor one in terms of explanatory value, suggesting little more than that the system is impaired from optimal efficiency, and delineates little but the conditions of its observation." Holland, therefore, prefers the first type of explanation, i.e. some change in the psychophysical scale, but it would take too long to go into the complex details of its discussion. For the purpose of this review we must conclude that the results, while strikingly large in terms of the small amount of drug administered, can only be explained in part by the hypothesis which the experiment was set up to test, and that further work will be required to test the alternative hypotheses advanced *ad hoc* to account for this failure of prediction.

8. *Flicker Perimetry*

In another experiment using the same subjects and drugs as the preceding one, Holland (Eysenck, 1960) tested the effect of depressant drugs on flicker at the limen of the visual field. The perimeter used was a modified Lister moving-arm type, which has its own illumination source for background and has targets of differing sizes and colors which are changeable. Modification to the instrument took the form of removing the normal target carried and mounting in its place a small neon bulb which subtended an angle of 1.8°. Flicker was provided by a pulse generator and the neon activated through a finger key. Owing to the modification the bulb carrier could not be positioned nearer than 4° from the center spot but from this position to 90° it was continuously variable along any meridian.

After some practice trials the critical flicker fusion (CFF) thresholds from flicker to fusion and fusion to flicker were assessed in fourteen different positions on the 45° and 135° meridians at 10°, 25°, 50°, and 60° peripheral angle, and along the 90° meridian at 10°, 30°, 45°, 60°, 70°, and 80° peripheral angle, all three coordinates being in the left temporal field. The method of limits was employed, and of the successive bursts of light none was longer than 3 seconds, with a 3-second interval between them.

Several analyses of variance were carried out all showing the same pattern, i.e. significant F ratios being derived for both in-

dividual differences and for treatment effects. None of the analyses are complicated by rank order effects. Both (M) and (D) depressed the CFF threshold, and graphs published by Holland (Eysenck, 1960) demonstrate that the depression is fairly uniform along the sensitivity gradient from the center to the periphery within the positions measured. This is in line with previous work on CFF which has usually demonstrated depression of CFF thresholds by depressant drugs and elevation by stimulant drugs.

9. *The Visual After-Image*

It is well known that stimulation of the eye by either a colored object or a rotating spiral produces certain after-effects once the stimulus is withdrawn. In the case of the colored object this is the well-known after-image; in the case of the rotating spiral it is an after-image which appears to produce effects opposite to those of the rotating spiral used as a stimulus. It is possible to relate the length of these after-effects to the drug postulate somewhat in this fashion. The experiment may be divided into two periods, (1) the period of *stimulation* during which the spiral or the colored surface is presented to the eye of the subject, and (2), after-effects or *reversal* period, during which the subject experiences a subjective sensation. In what follows I shall call the stimulation process X and the reversal $\overline{X}$; T refers to the time during which either process is in operation. During the period of stimulation (T) there is a constant increase in X; this may be assumed to take the form of a straight-line function, although the actual slope and shape of this line are irrelevant to the argument. This theoretical function, however, is interfered with by a process of *satiation* which reduces the amount of X to a degree which is proportional to the satiability of the subject, i.e. it will be greater for extraverts than for introverts, and for subjects after a depressant drug as compared with subjects after a stimulant drug. The total amount of stimulation actually received, therefore, is a function not only of T, but also of the state of the organism.

We now come to the reverse effect which is assumed to set in the moment the stimulus ceases to be effective, and which is assumed to be proportional in strength to the strength of the original stimulus effect, i.e. $\overline{X} = X$. Because of the original satiation, X, and

consequently $\overline{X}$, will differ from person to person in terms of their degree of extraversion, dose, and type of drug, etc. However, there is a further reason to predict that there will be differences in $\overline{X}$ between these groups. It must be assumed that the process underlying the reversal is physiological in nature and consequently subject to satiation; this satiation will again be stronger in extraverts than in introverts, and in subjects after a depressant drug as compared with subjects after a stimulant drug. These effects, leading to a shorter after-effect period for extraverts and depressant drug subjects as compared with introverts and stimulant drug subjects, are assumed to be cumulative; in any case the prediction is clear that after-effects will be shorter for extraverts and depressant drug subjects, and longer for introverts and stimulant drug subjects.

The evidence regarding personality correlates, particularly with respect to the spiral, indicates that the predicted relationship to extraversion-introversion is indeed found; both with normal and with neurotic subjects the predicted decrease in after-effect with extraversion has been observed (Eysenck, 1960). As regards drug effects, the only study available is one by Costello (1960b) in which he used a red square (5 cm^2) with a black focusing point in the center as a stimulus. This red square, surrounded by black opaque paper, was exposed in the viewing tube of a tachistoscope and illuminated by two 5-watt neon lamps behind a perspex sheet to which the red square was attached. All testing was done in a dark room. Six subjects were used in all and each subject was given two treatments: (P) and 600 mg (M). Both the (P) and (M) were in identically appearing tablet form and taken orally. They were administered on 2 different days, the order of treatments being counterbalanced. On each day the subject was given one session before treatment and four sessions 1, 2½, 3½, and 4½ hours after administration of the drug or (P). The red square was exposed for 30 seconds each time and after-effects noted by the subject on an automatic recording device. Two scores were used, one being the total duration of the after-effect, i.e., the time from the end of stimulation by the red square to the disappearance of the last after-image, the other the latency period, i.e., the time from the end of stimulation by the red square to the appearance of the first after-image.

Analyses of variance were performed and significant F ratios

emerged on both scores for subjects and treatments, and also for a subject/treatment interaction on the total duration scores. Costello concludes that the significant differences observed indicate (1) that considerable differences exist between people in the latency and total duration of the after-image, (2) that (M), as predicted, decreases the latency and total duration, and (3) that there are differences between people in the effect of (M) on the total duration of the after-image. It is noteworthy that although the subjects were tested at intervals up to 4½ hours after administration of the

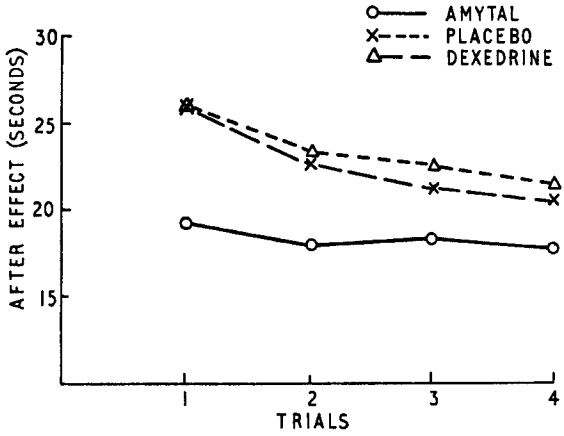

FIG. 9. After-effects of rotating spiral on subjects administered different drugs. From Eysenck *et al.* (1957c).

treatment and treatment/time interactions are not significant. The experiment on the whole decisively supports the hypothesis under investigation.

10. *Rotating Spiral After-Image*

After-effects of the rotating spiral were studied by Eysenck *et al.* (1957c). Six subjects were tested and were administered in counterbalanced order, in the form of a double Latin square, (P), 300 mg (A), and 10 mg (D-A). Four trials of 1-minute stimulation each, during two of which the spiral rotated in one direction, two in the other, constituted the testing procedure. Total length of after-effect provided the score. Results are shown in Fig. 9, and it will be seen that, as predicted, trials under (A) showed the shortest

after-effects, those under (D-A) the longest, with the (P) group intermediate. Analysis of variance showed highly significant F ratios for both people and drugs; the effects of replication were barely significant and none of the interactions were significant. Here again, then, results are positive and support the hypothesis under investigation.

A further study of the spiral after-effect duration was carried out by J. Easterbrook (unpublished data) the subjects being eight in number and the drugs being again 5 mg (D-A), 90 mg (A), 200 mg (M), and (P). The experimental design, a balanced incomplete block, insured that each drug would be given once after each other drug and in each serial position. The block was completed twice, once for the subjects seen in the morning and once for those seen in the afternoon. The mean duration of after-effects was 24.05 seconds under (P) conditions, 29.64 seconds under (D-A), 23.11 seconds under (A), and 24.97 seconds under (M). The over-all analysis of variance just failed to show a significant general effect of treatments, although for the most part these will be seen to have been in line with prediction. However, the greater duration of effects under the (D-A) as compared to the (P) treatment is significant by a t test.

11. *Suppression of the Primary Visual Stimulus*

The phenomenon used here is an objective after-image determination first introduced by Bidwell (1897). He found that if a brief stimulation of the eye with red light is followed by prolonged stimulation with white light, then what is seen is not a red flash but a green one. In the experimental arrangements used by Eysenck and Aiba (1957), a light source throws a shaft of white light through a red filter into a viewing tube; the subject holds one eye close to the other end of the viewing tube. The beam of light is interrupted by a rotating disc, into which a small sector has been cut; this sector allows a red beam of light to pass for a period of 20 msec before it is interrupted by the solid disc. The red beam is followed by a piece of white paper pasted on to the disc on the side of the subject and illuminated by a variable light source. Duration of exposure of the white stimulus was 125 msec. Starting with a very bright red, the subject's task is to reduce the intensity

of the stimulus until he sees only green and no red at all. Six subjects participated in the experiment, each serving as his own control under (P), (A), and (D-A) conditions. The prediction made was based on a theoretical explanation of this phenomenon. Briefly, the postulated chain of events is this. The brief red stimulus sets up two chains of events. One of these is photochemical, the other neural. Before the neural message reaches the cortex and is transformed into a red sensation, the white light is exposed and sets up preexcitatory inhibition (Granit, 1955) which destroys the neural effects of the red stimulus. Thus no red is seen, provided its brightness has not been too strong for the inhibitory impulse to overcome. At the same time that the red stimulation ceases, the photochemical activity is reversed and produces the green after-image. The image is now seen simultaneously with the direct sensation from the white stimulation, which thus produces a slight loss in saturation in the after-image. If (D-A) reduces the inhibitory effects postulated while (A) raises them, then we would expect the former to have the effect of lowering the threshold of the original red stimulus, while the latter would raise it. The predicted differences were in fact observed, all differences being significant at the 0.01 level or better, and the (P) group being intermediate between the other drug groups. The result therefore strongly supports the theory on the basis of which this test was selected.

A rather more complex experiment along similar lines was carried out by Aiba (unpublished results). Five subjects were used and tested before and after (D), 10 mg (D-A), and 300 mg (A). The procedure was as follows. After 15 minutes of dark adaptation the subjects were told to look at the stimulus patch and to increase the intensity of the red stimulus by turning the knob of the variac until they could see the first suggestion of red. As soon as they were satisfied with the adjustment, they resumed the dark adaptation until the next trial started. Four of these trials made up one session, and during the next 140 minutes the same procedure was repeated every 20 minutes. Thus there were altogether six sessions, each consisting of four trials. The drugs were administered between the second and the third session each day.

Results are shown in Fig. 10. Differences were shown to be significant at a very high level by means of analysis of variance, and it will be seen that as predicted the (D-A) lowered the threshold

whereas the (A) raised it. Of particular interest is the interconnection between drugs and sessions; "this showed that the drugs produced definite effects but that the time at which effects of each drug became manifest was not identical." This is an important point, because if the wrong time factor is chosen in experiments of this kind, duplication of results may be difficult.

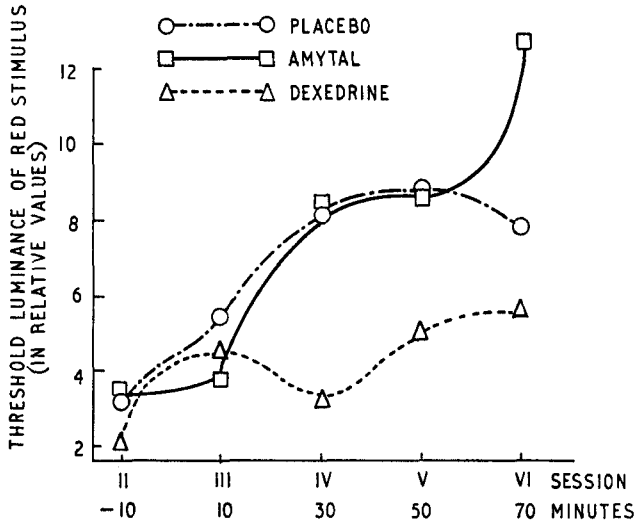

FIG. 10. Effects of drugs on the suppression of the primary visual stimulus, as a function of time since administration.

12. *Vigilance*

Vigilance is a term used for the behavior of test subjects who have to watch visual displays, or pay attention to auditory stimulation for lengthy periods at a time, always being ready to respond with a signal (pressing a bell, etc.) to certain relatively rarely occurring changes in the stimulus display. Thus, the subject may be required to watch a pointer going round a dial at a uniform speed; occasionally for a brief moment the pointer doubles its speed and this change in speed has to be noted (Mackworth, 1948).

On the theory under investigation one would expect that extraverts would be poorer at tasks of this kind (i.e., would tend to miss more signals, as compared with introverts) and similarly one would expect depressant drugs to cause subjects to miss more signals,

while stimulant drugs would be expected to have the opposite effect. The argument, of course, is quite a straightforward one depending on the accumulation of reactive inhibition in the neural structures responsible for what the layman calls "attention." There is ample evidence that the prediction regarding the relationship between vigilance and introversion is in fact as predicted, and there is some evidence from the work of Felsinger and his associates (1953) and of Mackworth (1948), that depressant drugs and stimulant drugs do indeed have the predicted effects.

Two experiments were carried out by E. Treadwell (1960) in an attempt to collect further evidence on this point. In the first experiment an auditory task was used. A series of single-digit numbers was recorded on tape and was reproduced at uniform loudness, the subject being instructed to react to any sequence of three odd numbers. The frequency with which the signals occurred was approximately six every 5 minutes, the time intervals between signals ranged from 3 seconds to 1 minute and 41 seconds with a mean interval of 58 seconds. The test lasted for 30 minutes and was preceded by a 2-minute period. Twenty-four subjects were tested after 400 mg (M), 250 mg (D), and no-drug conditions. Mean error scores for the whole test were 4.3 under no-drug conditions, 4.8 under (M), and 6.2 under (D). The results when tested by analysis of variance fell short of significance, although the mean scores are in the predicted sequence.

The test was found to be relatively poorly adapted to the needs of a study of this kind, and a new test was designed. In this, single-digit numbers were recorded on tape and played to the subject for 40 minutes, and he had to check these numbers against a printed list. Some of the numbers in the list were erroneous and these had to be indicated by the subject, and the correct number written in. Scores were calculated in two ways, one score consisting of the number of missed errors, the other one of the number of false responses. Eight subjects were tested on this test under conditions of (P), (D-A), and (A) administration.

The prediction that (D-A) would lower the number of missed and erroneous replies as compared with placebo conditions, while (A) would raise the number of missed and erroneous answers was realized. The number of missed responses for (D-A), (P), and (A) conditions, respectively, was 3, 7, and 14; the number of false

responses was 2, 13, and 18. Analysis of variance showed the "missed response" score to be statistically significant; the number of "false signal" scores did not quite achieve significance. On the whole, the results bear out the hypothesis and demonstrate that vigilance is impaired by depressant drugs and improved by stimulant drugs.

V. A Second Paradigm for the Testing of Psychopharmacological Hypotheses

The research design on which the investigations so far quoted have relied is illustrated in Fig. 11(*A*). It will be seen that this

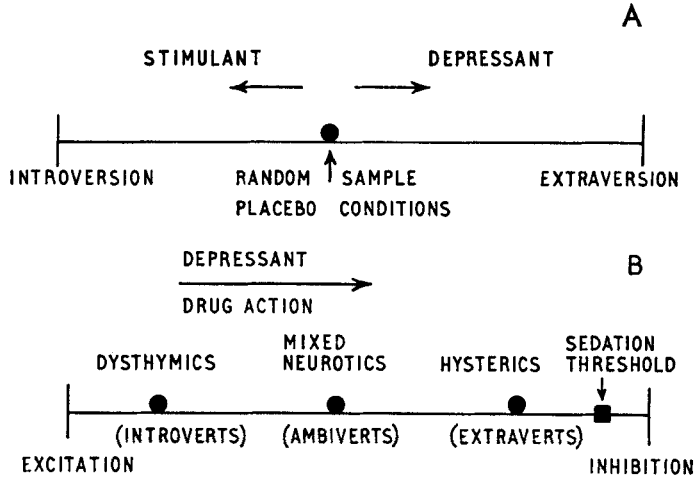

FIG. 11. Two fundamental research designs for the study of drug effects.

design does not depend in any way on an assessment of the personality of the subjects prior to the experiments. Subjects are allocated at random to control and experimental groups, or all subjects are tested under control and experimental conditions, and what is studied is a general effect of the drugs under investigation on what might be called the standard subject. This paradigm, is, of course, capable of certain improvements, but essentially it represents the great majority of research designs in present day psychopharmacology.

A rather different design is illustrated in Fig. 11(*B*). This design makes use of the known position of groups of subjects such as dysthymics and hysterics on the introversion/extraversion continuum, and thus explicitly contravenes the random sampling technique of design (*A*). This new design may be illustrated by going back to McDougall's observation that extraverts need less alcohol to reach the point of intoxication than do introverts, who, with the same amount of alcohol simply become more extraverted. Such a research design requires an objective *terminus ad quem*, i.e., the terminus of "intoxication threshold" which would enable us to ascertain the amount of alcohol required by the different groups of subjects to reach the same level of cortical inhibition as defined by this threshold.

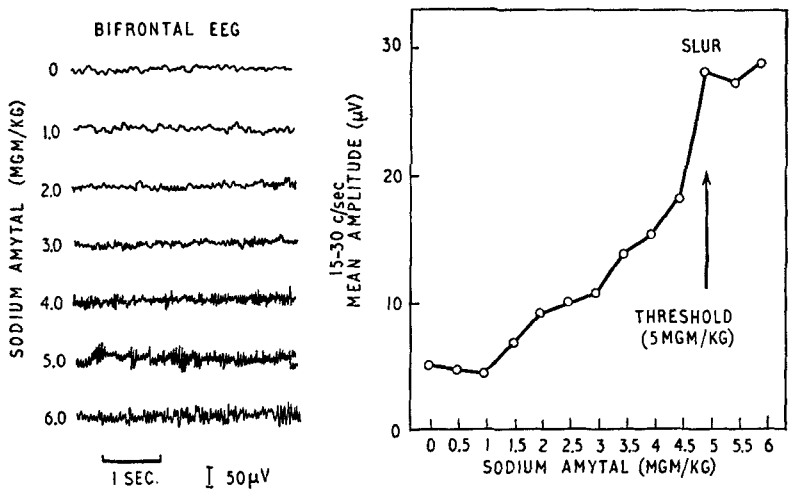

FIG. 12. Bifrontal EEG recordings illustrating sedation threshold.

The only example of the use of such a technique which the writer has been able to find is a study of Shagass (1954, 1956) and Shagass and Naiman (1955, 1956) using what he calls the "sedation threshold" of (A). The sedation threshold is an objective pharmacological determination, which depends on electroencephalogram (EEG) and speech changes produced by intravenous doses of (A) given at the rate of 0.5 mg/kg of body weight every 40 seconds. The patient is tested for slurred speech, and the injection is continued at least 80 seconds after slurred speech is noted. Con-

tinuous EEG's are recorded from transverse-frontal and sagittal-frontocentral placements. Figure 12 shows that (A) produces a rather striking increase of fast frequency (15–30 c/sec activity). The amplitude of this fast frequency is taken by Shagass as a response to the drug and the dosage-response curve plotted. "The typical curve has a sigmoid shape and contains a point of inflexion, preceding which there is a sudden increase in the amplitude of the fast activity, and following which the curve tends to plateau. This inflexion point generally occurs within 40 sec. (0.5 mg/kg) of the time when slurred speech is first noted, and the slur and inflexion point are used together as indicators of the threshold. The threshold is the amount of sodium amytal in mg/kg required to produce an inflexion point in the 15–30 c/sec. amplitude curve, which occurs within 80 sec. (1 mg/kg) of the time when the slur is noted. The slur localizes the threshold roughly, the EEG inflexion point does it more precisely. . . . The measurement is highly reliable; its probable error is not greater than 0.5 mg/kg of body weight. Age, sex, and previous intake of sedatives in usual psychiatric dosage have not been found to influence the threshold."

According to the theory outlined in this chapter (which was developed before Shagass's work was known to the writer), we should be able to make a very definite prediction. (A), being a depressant drug, would be postulated to increase inhibition. An extravert, whose cortex, according to our theory, is already in a relatively inhibited state, should require comparatively little (A) before reaching the critical sedation point; such a person should have a low sedation threshold. The introvert, on the other hand, whose cortex is in a state of considerable excitation and low inhibition, would require a considerable amount of (A) before reaching the critical sedation point; he would be predicted to have a high sedation threshold. If we express this general hypothesis in terms of neurotic groups and their standing on the extraversion-introversion continuum, then we would expect psychopaths to have the lowest threshold, followed by hysterics. Mixed neurotics would be intermediate and anxiety states, obsessionals, and reactive depressives would have sedation thresholds. An experiment along these lines was carried out by Shagass; his results are given in Fig. 13. It will be seen that these results bear out our prediction in every detail.

This design too, of course, is capable of certain interesting modifications. If something corresponding to the sedation or intoxication threshold could be found at the introverted end of the continuum, we would predict lower thresholds for dysthymics than for hysterics, and quite generally a reversal of the relations found

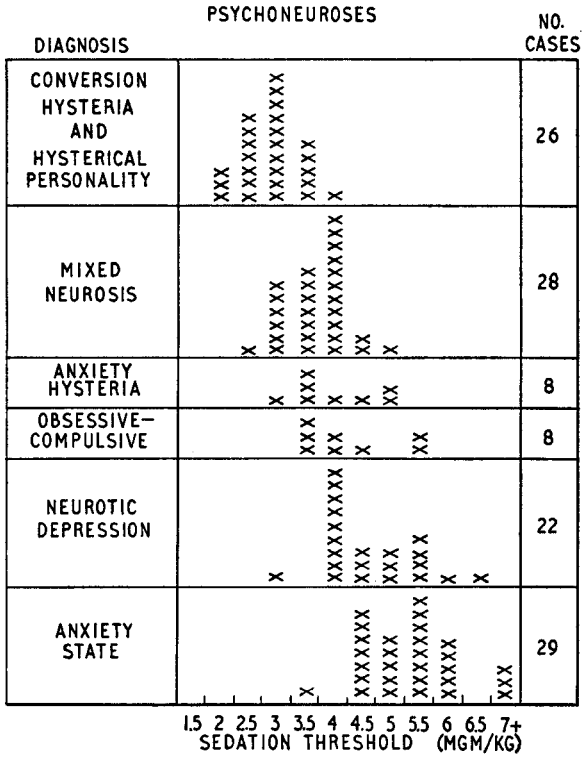

FIG. 13. Sedation threshold distribution for various neurotic groups.

on Shagass's research. Thus the same groups of patients at different times might be given different and opposing drugs as well as placebos. However, such developments depend on the discovery of such threshold effects for excitation drugs, which might possibly be looked for in EEG patterns corresponding to wakefulness as opposed to sleep.

VI. A Mathematical Model of a Psychopharmacological Hypothesis

The theory discussed in this chapter, it will be remembered, states that depressant drugs have an action on the cortex which increases inhibitory potential and decreases excitatory potential, while the stimulant drugs have an opposite effect. It will also be remembered that the classification of drugs presently offered by pharmacologists was merely accepted as a tentative first hypothesis, and that the ultimate criterion of the appropriateness of classifying a drug in one group or the other was to be its similarity of action on a group of behavioral tests to the action of other drugs in that group. Clearly these terms are too vague to be useful in a rigorous and systematic theory; we require a more clear-cut definition of such terms as "similarity." The method suggested here derives fundamentally from Fisher's "Discriminant function analysis," particularly that generalized form of it which goes by the name of "Canonical variate analysis." A detailed discussion of this has been given by Eysenck (1957b) and by Patrick Slater (1960), and consequently only a brief verbal description will here be attempted.

The first users of this method in psychology have all been concerned with a problem, which in some ways is rather similar to the one posed here. If we have groups of normals, neurotics, and psychotics, and if all these have been submitted to a battery of personality tests which discriminate adequately between them, can we use these scores to answer a very important psychiatric question, namely that of the relative position of these three groups? In other words, do normals, neurotics, and psychotics lie along one single dimension, with neurotics, as it were, less ill psychotics, and psychotics more seriously ill neurotics, or do we require two dimensions, i.e., one of neuroticism and one of psychoticism, to deal with the test performance of these subjects? What canonical variate analysis does in this situation is essentially this. It first of all combines all the scores obtained from our subjects in such a manner as to maximize the differences between the groups; in this way we obtain our first latent vector and its associated latent root. The method then goes on to extract another combination of scores uncorrelated with the first, and giving us the (residual) optimum differentiation between the three groups; this gives us the second latent vector and its associated latent root. If the three groups lie

along one dimension, then a test of significance applied to the second latent root would show it to be insignificant; in other words, all the differentiating power of the tests is accounted for in terms of the first vector. If, however, two dimensions are involved, then this is shown by the fact that the significance test reveals the second latent root to have a *P* value smaller than 0.05 (Eysenck, 1957b).

Let us now turn to the application of this principle to the model of a psychopharmacological theory. Let us suppose we are interested in a hypothesis to the effect that two drugs, say (M) and (D), are both CNS depressant drugs, i.e., that they differ from each other only in the strength of the effect produced, and not in the direction. We can test this hypothesis by experimentally creating three groups of subjects, i.e., those tested under (D), those tested under (M), and those tested under (P) or no-drug conditions. It would be possible to have different people in these three groups, and if tests were used scores on which changed considerably due to practice or learning, this would be the only experimental design open to us. However, in view of the great individual differences on test scores even under no-drug conditions, this design would be extremely wasteful and would require very large numbers of subjects; consequently a design seems preferable in which the three groups are made up of the same people tested under the three conditions in question.

To illustrate the method with an actual experiment, Eysenck and Eysenck (1960) selected five tests, namely: nonsense syllable learning, reaction times, level of skin resistance to the passage of an electric current, flicker fusion, and perimeter threshold difference; all these tests were administered to twenty-four subjects under conditions of no-drug, (D), and (M) medication. These five tests were chosen from a larger battery, the choice depending upon the total effectiveness of the test separating the groups, and on the necessity of maintaining experimental independance between the tests. Details of some of these tests have already been given on the previous pages. The results of the canonical variate analysis are shown in Table I. It will be seen that the first of the latent roots is significant at the 0.001 level, whereas the second latent root fails to be significant even at the 5% level. The results therefore bear out the hypothesis and show that both (M) and (D) are depressant drugs, and do not differ from each other in behavioral effects.

Scores were computed for each subject on both canonical variates in the following manner. The latent vectors gave two sets of weights to apply to the scores obtained from the tests, so that two measures could be calculated for each subject, one for each canonical variate. These scores, Y_1 and Y_2, were obtained by multiplying the score of the subject on the tests by the appropriate weights, and summing them over the five tests. The means of these scores are given in Table II. These scores enable us to

TABLE I

RESULTS OF CANONICAL VARIATE ANALYSIS

	Latent vectors	
Tests	X_1	X_2
Nonsense syllable learning	— 0.040,971	— 0.014,399
Reaction times	— 0.013,625	— 0.161,517
Level of skin resistance	— 0.080,835	— 0.013,827
Flicker fusion	1.000,000	1.000,000
Perimeter threshold difference	— 0.475,330	0.147,773

Latent roots

$\lambda_1 = 0.453,220 = 90.03\ \%$

$\lambda_2 = 0.050,205 = 9.97\ \%$

Diagonal entries of matrix $G^{-1}\ B = 0.503,425$

Significance of roots

$R_1^2 = \lambda_1,\ X^2 = 40.446,\ P < 0.001$

$R_2^2 = \lambda_2, X^2 = 3.450$, not significant

calculate what, by analogy with the psychiatric classifications problems, we may perhaps call the percentage of correct classification, i.e. given that we only knew the canonical variate scores of a person, could we correctly identify him as a person who had been tested under no-drug condition, under (M), or under (D)? The method used is one originally suggested by Maxwell (unpublished), and the results are shown in Table III; it will be seen that more than 70% of accurate classifications could be made under these conditions, which is a remarkable achievement considering (*a*) the very slight amount of drug administered, and (*b*) the fact that both drugs had the same effect, and were only differentiated from the point of view of the strength of the effect.

It may be useful to discuss briefly the consequences of the statistical results reported so far. In the first place, it is suggested that we have here the beginnings of a truly scientific system of classification of drugs. The experiment reported, of course, is only a very small beginning; only two drugs, five tests, and twenty-four subjects were employed. This, however, does not detract from the general value of the method itself, which could easily be

TABLE II

MEAN GROUP SCORES $\left(\overline{Y} = \frac{\Sigma W_1 Y}{n}\right)$

Subject	Means of scores
C	$\overline{Y_1} = 6.030$ $\overline{Y_2} = 8.622$
M	$\overline{Y_1} = 3.873$ $\overline{Y_2} = 9.047$
D	$\overline{Y_1} = 2.989$ $\overline{Y_2} = 8.485$

TABLE III

RESULTS OF CALCULATING THE PERCENTAGE OF CORRECT CLASSIFICATION[a]

	C	M	D
C	20	6	1
M	1	14	6
D	3	4	17
	24	24	24

[a] Under these conditions 70.8% correct classification could be made.

applied to much larger groups of drugs, subjects, and tests. In particular, it would be extremely interesting to include so-called stimulant as well as so-called depressant drugs to determine whether the unidimensionality of results could still be maintained. Furthermore, many other tests suggest themselves as being worthy of inclusion, and just as we can assess classificatory properties of the drug in terms of the test results, so we would then be able to assess the psychological significance of each test in terms of the displacement of scores on the test by sets of drugs found to be homogeneous in their action. In the particular case of stimulant

and depressant drugs, we would also be able to test the classification of drugs in terms of tests by comparing the results with those of testing typically extraverted and introverted groups. In this way we would treat change in test scores as the dependent variable, and drugs on the one hand, and selection of extraverted or introverted subjects on the other, as the independent variable.

It is necessary, however, to insert one word of caution at this point. If a latent root derived from the analysis of a given set of test scores is significant, then it is clear that the dimensionality of the universe in question cannot be less then that indicated by the number of significant roots, although it may be greater. (Thus in the psychiatric example mentioned earlier, another dimension additional to that of neuroticism and psychoticism, e.g. extraversion-introversion, might be involved, but could not be discovered by statistical analysis unless the groups were further subdivided, e.g. into hysterics and anxiety states, etc.) In this way the canonical variate analysis is inferior to factor analysis by enabling us only to test the significance of specific hypotheses which are embodied in the choice of groups and tests; factor analysis, although it has other shortcomings, is not so restricted.

When, however, one or more of the latent roots are found to be insignificant, it does not necessarily follow that no significant differences exist between groups. A faulty choice of tests may give the impression of lack of difference, when a suitable choice of tests might have given quite different results. There is no possibility of guarding against this, and particularly at the beginning of the study of such a vast field as this it is almost certain that choice of tests will be relatively haphazard, and erroneous conclusions may easily be drawn. In the present case, for instance, it is possible that if reports of subjective feelings had been included as one of the tests, then a significant second root might have been established. (Subjects taking (M) reported no after-effects, while those taking (D) reported feelings of sleepiness, etc.)

The solution to the problem is twofold. In the first place, growth of knowledge and increase in the number of properly conceived and executed studies in this field will undoubtedly soon help us to cut down drastically the numbers of errors of this kind, and will help us to design batteries of tests which will cover the whole field more completely than is possible at present.

The second solution to the problem is complementary to the first. What is required more than anything, in our opinion, is the creation of a firm set of theories and hypotheses linking psychopharmacology with general and experimental psychology. It is by testing, confirming, revising, and improving theories of this kind that we must hope that improvement in the present unsatisfactory situation will ultimately come. Such has been the development of most scientific subjects, and it is difficult to see why psychology should be an exception to this rule.

VII. The Place of Animal Work in a Hypothetico-deductive System

Animal experiments play a large and important part in the study of drugs, but it has always been axiomatic that results from animal work cannot be accepted indiscriminately as applying to human beings as well. Indeed, there has been much argument in the psychopharmacological field as regards the propriety of accepting results from animal work as being in any way relevant to human behavior. While this chapter is concerned in the main with the testing of pharmacological hypotheses in human subjects, it may be relevant to discuss quite briefly the place of animal experimentation within the hypothetico-deductive system under discussion.

The difficulties of interpretation of animal work derive precisely from the same sources as do the difficulties of interpretation of work with human subjects discussed in the first section. Behavior, and changes in behavior, can only become meaningful as part of a theoretical system. If that system makes predictions in the animal field, then animal experimentation becomes relevant. The system under discussion here is concerned with such concepts as excitation and inhibition in the central nervous system, and in terms of modern learning theory, these concepts are just as relevant to animal behavior as they are to human behavior; indeed, some critics might claim that they were more relevant to animals than humans! Consequently we may use animals to confirm or refute specific hypotheses, and in turn animal work may give us helpful suggestions as to experiments to be performed with human subjects, and tests to be used in the human field.

As this is a relatively minor point, only one example will be given of our work in the animal field. The study itself is concerned with alternation behavior in a simple T-maze (Sinha *et al.*, 1958). Such alternation behavior—that is, turning in the direction opposite to that of previous turns—is conceived of as a consequence of reactive inhibition; in other words, turning right at one point sets up reactive inhibition to further right turns and predisposes the animal to turn left, and vice versa. It is known that the amount of alternation is a function of the number of preceding turns, and it can be

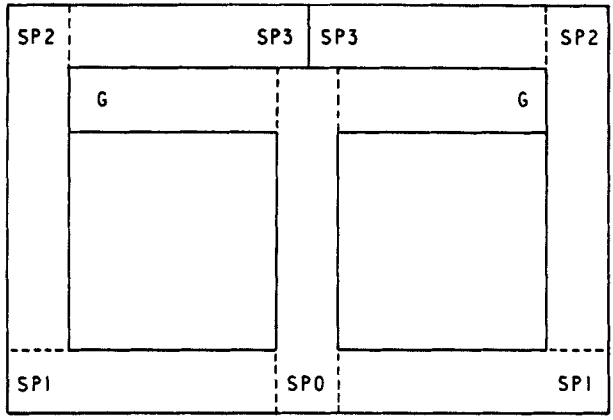

FIG. 14. T-maze used in experiment on drugs as affecting alternation behavior. Taken from Sinha *et al.* (1958).

predicted from the general drug postulate that this function will be increased by a depressant drug and decreased by a stimulant drug.

The floor plan of the maze used is shown in Fig. 14. It consists of a rectangular runway communicating with a T-maze only at the lower end of the upright of the T. The alley forming the other side of the rectangular runway, opposite this junction, was permanently divided into two halves. Thus by suitable manipulation of a series of temporary blocks within this outer runway, subjects could be run to a food reward either from the foot of the T (starting position 0; SP0) or from either of the lower corners of the outer rectangle, left or right (starting position 1; SP1), or from either of the upper corners (starting position 2; SP2), or from either side of the permanent block on the upper side of the rectangle (starting posi-

tion 3; SP3). An animal would thus have made 0, 1, 2, or 3 turns, respectively, before entering the final T and these turns could be either to the right or to the left.

Sixty-four rats in all were run through the maze, the drugs used being (A), 15 mg/kg of body weight, (P), and Meratran, 10 mg/kg of body weight. Each animal was tested under each of the three conditions. All rats were given preliminary adaptation training on the apparatus, and throughout the training they were forced to one or the other of the goal arms of the T by blocking the opposite side at the junction according to a Gellerman series, in order to guard against the development of position habits. An analysis showed that the training procedures, designed to establish equal asymptotic habit strength in running to either right or left goal from the various starting positions, were in general successful.

All subjects were then given the same alternation test lasting two days, which consisted in allowing the rat a free choice in the T, both goal arms now being left open. This was done on eight trials on two successive days, and each such free-choice trial was followed by a forced trial with one or other of the arms blocked. Thus, a total of eight trials, four free alternating with four forced, was given each day with a constant intertrial interval, despite the differences in starting position. The starting position side was selected as before—for rats other than those in group 0—by use of a modified Gellerman series, but the goal side which was blocked on the forced trials was in each case determined by rat's own choice on the previous free-choice trail, so that each rat was forced away from the side previously chosen. This was a further precaution against the development of position habits.[3]

During each free-choice trial an alternation response was scored only when the rat turned to the goal arm of the T in the direction opposite to that of its previous turns in the maze on that same trial. A rat with a position habit would thus not vitiate the results because it would, by always turning to one side and being started from each side four times, score only four out of the possible eight alternations—that is, chance level. Alternation responses for rats from starting position 0, without previous turns, were scored in a comparable way, the justification for which is the assumption that

[3] Data in this paragraph from Sinha *et al.* (1958).

I_R may persist from one trial to the next in this simple situation (Zeaman and House, 1951). In this case, an alternation response was defined as a turn on free-choice trial to the goal arm of the T in the direction opposite to that to which it was forced on the previous trial. For the first (free-choice) trial on any day, the scoring convention adopted was to regard the last (forced) trial on the previous day as influential. The possibility of artifacts in

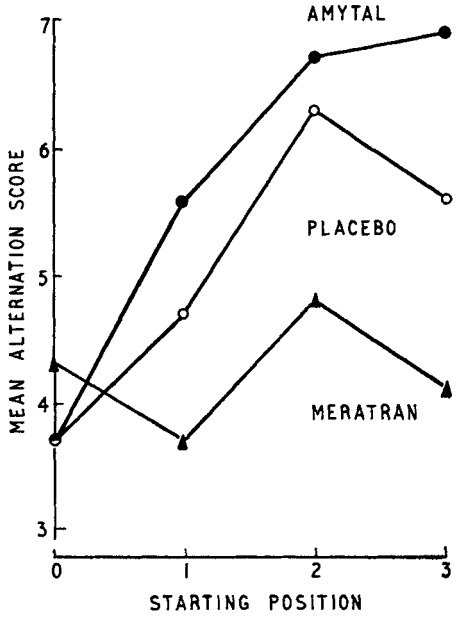

FIG. 15. Effect of drugs on alternation behavior in rats. Taken from Sinha *et al.* (1958).

the data caused by position habits among these rats is more serious as it is not possible to reduce their effect to chance level in the manner described above. Moreover, by our scoring, a rat which always turned to the same side would receive the maximum alternation score of eight, since it would in each case be forced to the opposite side on the intermediate trials. Examination of the data, however, showed that no rat in this group responded consistently in this way.[4]

Results are shown in Fig. 15. It will be seen that in general

[4] Data in this paragraph from Sinha *et al.* (1958).

there is an increase in alternation behavior as a function of the number of preceding turns, and it will also be seen that alternation behavior is increased under (A) and abolished under Meratran. Analysis of variance showed these effects to be highly significant. It may be concluded that the prediction of the drug postulate is verified.

Many fairly obvious experiments suggest themselves for use with human subjects as a consequence of the demonstration that alternation behavior is predictably affected by stimulant and depressant drugs. Almost any situation in which a choice has to be made between two or more stimuli, and in which such choice is repeated fairly regularly, might be used in this connection. The choice might be physical, as in the choice of one of different paths, or in the choice of one of different foods at meal time; or it might be perceptual, as in the choice of one of several pictures to look at, or one of several films to see. There has been a good deal of controversy about the perceptual and motor theories of alternation behavior (House, 1956), but it would seem that the generalization mediating our prediction would equally well embrace both types of behavior.

VIII. A General Dimensional System of Psychopharmacology

In the preceding sections, an attempt has been made to set forth and to illustrate one particular hypothesis of psychopharmacological action, and to work out the paradigms appropriate to the use of the hypothetico-deductive method in this context. It should not be thought, however, that the general principles and methods here suggested are restricted to the dimension of extraversion-introversion, or the effects of stimulant and depressant drugs. The method is a perfectly general one, and it applies in principle to all dimensions of personality (Eysenck, 1947), and to those groups of drugs affecting a person's position on these dimensions. The only limitation at present appears to be that a proper dimensional analysis of the relevant personality traits must have been carried out before work in the psychopharmacological field can be commenced; this of necessity means that only a relatively small number of dimensions can at present be studied in this manner (Eysenck, 1959d).

Apart from intelligence and the general field of cognitive abilities, the only two dimensions which have been identified sufficiently well to furnish the necessary substratum for work of this kind are *neuroticism* or general emotional instability, and *psychoticism.* The evidence regarding these factors has been discussed elsewhere by the writer (Eysenck, 1959d, 1960), and we need only state here that these two factors appear to be orthogonal to each other and to extraversion-introversion (Eysenck, 1957b; Eysenck, S. B. G., 1956); intellectual ability is largely independent of all three personality factors.

It is likely that drugs which stimulate the sympathetic division of the autonomic nervous system, and which inhibit the parasympathetic division, increase neuroticism, while drugs which inhibit the sympathetic division and stimulate the parasympathetic division are likely to decrease neuroticism (Trouton and Eysenck, 1960); there is too little systematic evidence to make it possible at the moment to arrive at any clear answer. The possibility does exist, however, of finding such an answer by applying the methods outlined in this chapter to the large group of tests known to measure neuroticism (Eysenck, 1959d; Eysenck, 1952b; Eysenck *et al.*, 1957b). An attempt to carry out such a project would be likely to advance our knowledge of both neuroticism and of drug action to a considerable extent.

As regards psychoticism, the position is even less satisfactory than with respect to neuroticism. A discussion can perhaps begin by quoting some comments made by Freyhan (1959), apropos of suggestions made by various speakers at a psychopharmacological conference, that certain drugs had "antischizophrenic and/or antipsychotic effects." Freyhan said: "I question the legitimacy of these terms. It does not make clinical sense to say 'antischizophrenic' when schizophrenic must include great varieties of syndromes such as muteness, explosive aggressiveness, grandiose omnipotence or cosmic despair. Are we to assume that all these typical syndromes are to be acted upon by the same drug in the same manner with the same effectiveness? And I wonder about the applicability of the formulation 'more psychotic' in reference to a patient who, under the influence of a stimulant drug, changed from inconspicuous to disturbed or disturbing behavior. Was this patient 'less psychotic' when he sat quietly in a corner, surrendering to bizarre fantasies

which he kept to himself? Neither psychosis nor schizophrenia denote molecular units which can be quantitatively analyzed or, indeed, be construed as therapeutic targets per se."

While recognizing the truth of the criticism as applied to much current work, the conclusion expressed in the final sentence cannot be supported by available evidence. The question of whether there is one particular general factor running through all psychotic disorders is not one which permits of being answered by means of simple clinical observation; it requires an answer in terms of properly planned experiments. Such evidence as there is suggests that there does exist a continuum from normality to psychotic disorder, and that this may properly be called psychoticism; it appears furthermore that objective measures of this continuum can be constructed and used with sufficient precision to make possible the assessment of "antipsychotic" effects of any particular drug (Eysenck, 1952a,b). It also appears that there does exist some functional unity in the concept of schizophrenia, apart from the general notion of psychoticism, and that this also is measurable by means of objective tests (Payne and Hewlett, 1960). While these results are by no means definitive, they do suggest that experiments aimed at the demonstration of antipsychotic or antischizophrenic effects of certain drugs may be by no means premature.

Let us apply these considerations to another point made by Freyhan. He maintains that "drug therapy must aim at target symptoms. These symptoms, regardless of diagnostic entities, can be fairly specific and are often mutually exclusive." This statement is true in its second part, but leaves out the important fact that each patient has a position not only on one dimension (psychoticism, say) but also on all other dimensions of personality; it is his position on these other dimensions which may determine the visible symptom. However, drug treatment is concerned with changes in the underlying causes of his psychoticism; the specific symptoms are merely incidental and may be presumed to disappear once the correct general antipsychotic drug has been applied. Such considerations would of course not apply to drugs which are only meant to treat the symptom itself; they might "cure" specific symptoms in one psychotic patient and exacerbate them in another. This would indicate that we were dealing with a drug affecting a dimension of personality orthogonal to psychoticism, but interacting with it in producing the symptom in question.

These points have been made in some detail because they are vital to a proper understanding of the consequences of introducing the dimensional system of personality description into the psychopharmacological field. Symptoms are often determined by several dimensions at once; they are therefore subject to all the difficulties discussed in the second section of this paper. Anxiety is perhaps the outstanding example of such a symptom; it has loadings both on the factors of neuroticism and of introversion (Eysenck, 1957b, 1959d). Changes in anxiety due to the administration of a drug cannot therefore be interpreted in themselves; the drug may affect introversion, neuroticism, or both. A factorial design alone can give us an answer to this question.

IX. Summary

In this review the writer has tried to summarize recent work carried out on a specific theory of psychopharmacological action, and to indicate how the methodology developed in order to deal with the problems arising in this attempt could be generalized to solve similar problems in other areas.

It was pointed out that most present-day studies, however perfect they might be technically, yet fall short of making possible a psychological interpretation of the results because they do not take into account (*a*) the multifactorial nature of the score on any given test, or (*b*) the determination of performance by several different and distinct factors (drive, habit, inhibition, etc.). It was then pointed out that these difficulties could only be overcome by means of the formulation of theories and hypotheses linking drug action with established personality dimensions, and by the use of certain advanced statistical techniques (factor analysis, discriminant function analysis, canonical variate analysis.)

A specific hypothesis was stated, to the effect that depressant drugs increase cortical inhibition and decrease cortical excitation, while stimulant drugs have an opposite effect. Twelve deductions from this hypothesis were made, and experiments confirming or infirming these deductions discussed. It was concluded that on the whole the results were in line with the hypothesis. Results from an alternative experimental paradigm were also found to support the hypothesis.

In the final section, an attempt was made to extend the general approach to other dimensions of personality, and to show that psychiatric symptoms require to be dealt with in the same manner as do scores on psychological tests, and that they are subject to the same difficulties and complexities. It was concluded that only by introducing the concepts of dimensional analysis from the field of personality research into psychopharmacology could the many discrepancies and contradictions so rife at present be reconciled.

References

Bidwell, S. (1897). *Proc. Roy. Soc.*, **B61**, 268.

Burstein, E., and Dorfmann, D. (1959). *J. Psychol. (London)* **47**, 81.

Costello, C. G. (1960a). *J. Mental Sci.* 322.

Costello, C. G. (1960b). *J. Mental Sci.* 326.

Eysenck, H. J. (1947). "Dimensions of Personality." Routledge & Kegan Paul, London.

Eysenck, H. J. (1952a). *J. Pers.* **20**, 345.

Eysenck, H. J. (1952b). "The Scientific Study of Personality." Routledge and Kegan Paul, London.

Eysenck, H. J. (1957a). *J. Mental Sci.* **103**, 119.

Eysenck, H. J. (1957b). "Dynamics of Anxiety and Hysteria." Routledge & Kegan Paul, London.

Eysenck, H. J. (1958). Personality tests: 1950-55. *In* "Recent Progress in Psychiatry" (Fleming, G. W. T. H., ed.), pp. 118-159. J. A. Churchill, London.

Eysenck, H. J. (1959a). "Manual, Maudsley Personality Inventory." Univ. of London Press, London.

Eysenck, H. J. (1959b). "Behaviour Therapy and the Neuroses." Pergamon Press, New York.

Eysenck, H. J. (1959c). *In* "Drugs and Behavior" (J. G. Miller, ed.). Wiley, New York. In Press.

Eysenck, H. J. (1959d). "The Structure of Human Personality" (2nd ed.). Methuen, London.

Eysenck, H. J. (1959e). *Brit. J. Psychol.* 360.

Eysenck, H. J. (ed.). (1960). "Experiments in Personality." Routledge & Kegan Paul, London.

Eysenck, H. J. and Aiba, S. (1957). *J. Mental Sci.* **103**, 661.

Eysenck, H. J., and Eysenck, S. B. G. (1960). *In* "Experiments in Personality" (H. J. Eysenck, ed.). Routledge & Kegan Paul, London.

Eysenck, H. J., Casey, S., and Trouton, D. S. (1957a). *J. Mental Sci.* **103**, 645.

Eysenck, H. J., Granger, G., and Brengelmann, J. (1957b). "Perceptual Processes and Mental Illness." Chapman & Hall, London.

Eysenck, H. J., Holland, H. and Trouton, D. S. (1957c). *J. Mental Sci.* **103**, 650.

Eysenck, S. B. G. (1956). *J. Mental Sci.* **102**, 517.

Felsinger, J. H., Lasagna, L., and Beecher, H. K. (1953). *J. Pharmacol. Exptl. Therap.* **109**, 284.
Fleishman, E. A. (1956). *Air Force Human Eng., Personnel, Training Research, Tech. Rept.* **56-8**, 142.
Fleishman, E. A., and Hempel, W. E. (1954). *Psychometrika*, **19**, 239.
Fleishman, E. A., and Hempel, W. E. (1955). *J. Exptl. Psychol.* **49**, 301.
Franks, C. M., and Laverty, S. G. (1955). *J. Mental Sci.* **101**, 654.
Franks, C. M., and Trouton, D. (1958). *J. Compand Physiol. Psychol.* **51**, 220.
Freyhan, F. A. (1959). "Biological Psychiatry," p. 323. Grune and Stratton, New York.
Granit, R. (1955). "Receptors and Sensory Perception." Yale Univ. Press, New Haven, Connecticut.
Hauty, G. T., and Payne, R. B. (1955). *J. Exptl. Psychol.* **49**, 60.
Holland, H. (1960). *In* "Experiments in Personality" (H. J. Eysenck, ed.). Routledge and Kegan Paul, London.
House, J. B. (1956). *J. Genet. Psychol.* **89**, 183.
Hull, C. L., Hovland, C. I., Ross, R. T., Hall, M., Perkins, D. T., and Fitch, F. B. (1940). "Mathematico-deductive Theory of Rote Learning." Yale Univ. Press, New Haven, Connecticut.
Laverty, S. G., and Franks, C. M. (1956). *J. Neurol. Neurosurg. Psychiat.* **19**, 137.
McEwen, P. (1958). *Brit. J. Psychol. Monogr. Suppl.* **31**, 1-92.
Mackworth, H. (1948). *M.R.C. Special Report No.* **268**.
Martin, I. (1960). *In* "Experiments in Personality." (H. J. Eysenck, ed.), Routledge & Kegan Paul, London.
Payne, R. B., and Hauty, G. T. (1953). *U.S.A.F. Sch. Aviat. Med. Progr. Rept. No.* **21**.
Payne, R. B. and Hauty, G. T. (1955). *J. Exptl. Psychol.* **50**, 343.
Payne, R. W., and Hewlett, J. H. G. (1960). *In* "Experiments in Personality." (H. J. Eysenck, ed.), Routledge & Kegan Paul, London.
Poser, E. G. (1958). *Am. Psychol.* **13**, 334.
Shagass, C. (1954). *Electroencephalog. and Clin. Neurophysiol.* **6**, 221.
Shagass, C. (1956). *Psychosomat. Med.* **18**, 410.
Shagass, C., and Naiman, J. (1955). *A.M.A. Arch. Neurol. Psychiat.* **34**, 397.
Shagass, C., and Naiman, J. (1956). *J. Psychosom. Research* **1**, 49.
Sinha, S. N., Franks, C. M., and Broadhurst, P. L. (1958). *J. Exptl. Psychol.* **56**, 349.
Slater, P. (1960). "Experiments in Personality" (H. J. Eysenck, ed.). Routledge & Kegan Paul, London.
Summerfield, A., and Steinberg, H. (1957). *Quart. J. Psychol.* **9**, 146.
Taylor, J. (1951). *J. Exptl. Psychol.* **41**, 81.
Treadwell, E. (1960). *In* "Experiments in Personality" (H. J. Eysenck, ed.). Routledge & Kegan Paul, London.
Trouton, D. S., and Eysenck, H. J. (1960). *In* "Handbook of Abnormal Psychology" (H. J. Eysenck, ed.), Pitmans, London.
Willett, R. A. (1960). *In* "Experiments in Personality" (H. J. Eysenck, ed.), Routledge & Kegan Paul, London.
Zeaman, D. and House, B. J. (1951). *J. Exptl. Psychol.* **41**, 177.