

Mischel and the concept of personality

Michael W. Eysenck and Hans J. Eysenck

The various criticisms that Mischel has made of the state–trait approach to personality are considered and found to be lacking in substance. His major argument is that the actual inconsistency of behaviour is incompatible with the expectation of behavioural consistency that follows from the state–trait approach. However, Mischel has misread the evidence, and pays insufficient attention to the distinction between consistency at the intervening-variable level and consistency at the behavioural level. In addition, Mischel and others have evaluated state–trait theories from a rather narrow perspective and thus have failed to appreciate the substantial contribution made by such theories. It is concluded that personality forms an indispensable part of experimental and applied psychology, and that Mischel's criticisms have unfortunately tended to accentuate the schism between personality and experimental psychologists.

Over the past decade, there has been increasing criticism of the state–trait approach to personality. While doubts had been expressed previously, for example by Vernon (1964), it was the publication of a book by Mischel (1968) that provided the impetus for much of the subsequent debate. For purposes of expository clarity, it will be assumed that state–trait theorists (e.g. R. B. Cattell, H. J. Eysenck, J. P. Guilford) share the following preconceptions about the most appropriate approach to theorizing in the field of personality:

- (1) Individuals differ with respect to their location on important semi-permanent personality dispositions, known as 'traits'.
- (2) Personality traits can be identified by means of correlational (factor-analytic) studies.
- (3) Personality traits are importantly determined by hereditary factors.
- (4) Personality traits are measurable by means of questionnaire data.
- (5) The interactive influence of traits and situations produces transient internal conditions, known as 'states'.
- (6) Personality states are measurable by means of questionnaire data.
- (7) Traits and states are intervening variables or mediating variables that are useful in explaining individual differences in behaviour to the extent that they are incorporated into an appropriate theoretical framework.
- (8) The relationship between traits or states and behaviour is typically indirect, being affected or 'moderated' by the interactions that exist among traits, states, and other salient factors.

The Thorndike–Mischel critique: Behavioural consistency

Theories of this kind, be they trait or type theories, have been most forcefully criticized by Thorndike (1903), who held that 'there are no broad, general traits of personality, no general and consistent forms of conduct which, if they existed, would make for consistency of behaviour and stability of personality, but only independent and specific stimulus–response bonds or habits' (p. 29).^{*} This doctrine of 'Sarbondism', as McDougall used to

^{*} Other typical statements of early situationism from Thorndike (1903) are the following: 'The striking thing is the comparative independence of different mental functions even where to the abstract psychological thinker they have seemed nearly identical. There are no few elemental faculties or powers which pervade each a great number of mental traits so as to relate them closely together' (p. 28). And again: 'The mind must be regarded not as a functional unit nor even as a collection of a few general faculties which work irrespective of particular material, but rather as a multitude of functions each of which is related closely to only a few of its fellows, to others with greater and greater degrees of remoteness and to many to so slight a degree as to elude measurement' (p. 29).

refer to it, with its attending notion of the equipotentiality of the CS, has by now more or less disappeared from psychology, and does not therefore require an extended answer; it may, however, be useful to point out that even within 'Sarbondism' consistency of behaviour and personality is by no means ruled out. It is not difficult to envisage conditions of life which would favour the production of consistent sets of S-R bonds which might give rise to certain traits; thus soldiers in a Guards regiment would be subjected to many conditions in which tidiness would be rewarded, and untidiness punished. This should, even on Thorndike's own grounds, give rise to consistently tidy behaviour, or a trait of 'tidiness'. In more modern terms, a consistent history of reinforcement should be able to create consistent forms of behaviour, and persistent behavioural traits and types.

This division of opinion regarding consistency of conduct gave rise to many experiments in the 20s and 30s; these have been reviewed by H. J. Eysenck (1970) in some detail. He concluded that these studies gave unambiguous evidence of consistency of behaviour, even when, as in the case of the large-scale work on honesty, deceit, self-control and organization of character (Hartshorne & May, 1928, 1929; Hartshorne & Shuttlesworth, 1930), the original authors drew an opposite conclusion from their data. H. J. Eysenck also discussed in detail the applicability of many of the criticisms later made of the concept of consistency, and showed them to be largely mistaken.

More recently, Mischel (1969) has taken up the argument, suggesting that while trait theory predicts behavioural consistency, it is behavioural inconsistency that is typically observed. He writes: 'I am more and more convinced, however, hopefully by data as well as on theoretical grounds, that the observed inconsistency so regularly found in studies of noncognitive personality dimensions often reflects the state of nature and not merely the noise of measurement' (p. 1014). The basis for this assertion was the partial review of the relevant literature by Mischel (1968), who concluded that measures of consistency in personality rarely produce correlations as high as 0.30.

Mischel's argument is subject to the same criticisms as Thorndike's, and these will now be presented very briefly; a more extended discussion of the evidence supporting these criticisms, together with a review of much of the empirical evidence, is presented by H. J. Eysenck (1970). We should note, however, that while Thorndike wrote at a time when the evidence was ambiguous, and too fragmentary to allow of any certain conclusions as regards the consistency of conduct, the evidence is by now so voluminous, and so strong and unambiguous, that it is curious that Mischel's doctrines should have attracted as much attention as they have. Boring would no doubt have explained this fact by appealing to the *Zeitgeist*, which floats like a disembodied spirit above his *History of Experimental Psychology*; we put forward no hypothesis in this connection.

At the empirical level, an inadequacy of many studies has been the use of very limited and unreliable data sampling. The difference that enlarging the data base can make to correlational measures of consistency was demonstrated clearly by Epstein (1977). Subjects kept records of their most positive and negative emotional experience each day for over 3 weeks. The mean correlation when either positive or negative experiences were compared on only 2 days was less than +0.20, and very much in line with the magnitude of most of the correlations discussed by Mischel (1968). However, when the mean for all the odd days was correlated with the mean for all the even days across subjects, the mean correlation for the pleasant emotions was +0.88, and was only slightly less for the unpleasant emotions.

The above findings are, of course, based entirely on self-report data. However, Epstein (1977) also discussed observations made daily by external judges for 4 weeks on eight variables related to sociability and impulsivity. The mean correlation based on two 1-day samples of behaviour was +0.37, versus +0.81 for two 14-day samples, and the highest reliability coefficients were produced by those variables requiring the least inference.

One of the problematical aspects of the Mischel critique is that he sometimes seems to imply that the putative consistency of personality can be effectively discredited by reference to the situational specificity of behaviour. For example, Mischel (1973*b*) argued that, 'People may proceed quickly beyond the observation of some consistency which does exist in behaviour to the attribution of greater perceived consistencies which they construct' (pp. 341–342). The implication that the only place to look for consistency is in overt behaviour is surely erroneous. Since both trait and state concepts are intervening variables, one must distinguish between consistency at the mediating level of states and traits, and consistency at the level of specific behavioural responses. It would be unreasonable to deny the possibility that specific behavioural inconsistency may coexist with a more conspicuous consistency at the mediating level.

In essence, the data suggest that reasonably high consistency at the intervening-variable level is accompanied by apparently inconsistent and situation-specific behaviour. Block (1977) evaluated the three main kinds of personality data: objective test behaviour, self-report, and rating. He concluded that self-report and rating data are often reliable and also comparable, but that objective test data tend to be unreliable and inconsistent. Mischel's evidence of low reliability coefficients centred, of course, on objective test responses. Even here recent studies, and the proper evaluation of earlier studies such as those of Hartshorne & May, give evidence of impressive consistency.

Mischel's criticism leaves out of account the simple fact that complex traits (e.g. 'honesty') cannot meaningfully be measured by a single, simple behavioural test. As the Hartshorne & May studies have shown, intercorrelations between such simple tests are only +0.2 or thereabouts, giving negligible prediction of actual behaviour as rated by teachers; when a battery of nine such behavioural tests is used, however, it has considerable reliability, and correlations with outside, real-life criteria are between +0.5 and +0.6. Thus even behavioural data, when properly used, can give strong evidence of consistency; inappropriate usage, of course, should not be accepted as evidence against consistency.

It is interesting that Mischel (1977) has now accepted that ratings by observers and self-ratings can both show impressive reliability and consistency over time. However, the proper interpretation of these findings is in dispute. Mischel (1968, 1977) argued that the perception of personal consistency in ourselves and others involved the imposition of order, and that this served the function of reducing the otherwise unmanageable complexity of the actual situational specificity of behaviour. Mischel (1968) expressed the argument in the following way: 'The conviction that highly generalized traits do exist may reflect in part (but not entirely) behavioural consistencies that are *constructed* by observers, rather than actual consistency in the subject's behaviour' (p. 43). Finally, Mischel implied that the observation of actual behaviour provided the basis for an objective approach to the study of personality.

Mischel's critique: Some counter-arguments

(1) One of the best-known of Mischel's criticisms of the state-trait approach is his assertion that measures of consistency in personality rarely produce correlations in excess of +0.30. This criticism is applicable at most to studies considering specific behavioural responses across two dissimilar situations. As we have seen in the work of Epstein (1977), reliability coefficients greater than +0.80 can be obtained in self-report and rating data.

(2) Mischel has frequently argued that traits are constructs which are inferred from behaviour, implying that the concrete behaviour which is observed is somehow objective. It must be doubted whether any straightforward distinction between the objective nature of behavioural facts and the subjective way we interpret them is justified. Experimenters

invariably use implicit or explicit theoretical notions to define the particular response-equivalence classes that are to be used in data collection. For example, Skinner (1938) constructed a single-response class, with all responses of sufficient strength to depress the lever being considered as equivalent, and all other responses being ignored. It is a matter of opinion whether the theoretically based selectivity of observation and utilization of a limited number of arbitrarily chosen response-equivalence classes should be construed as objective in any important sense.

(3) The issue concerning response classes is also relevant to Mischel's position in a rather different way. It is a plausible assumption that individuals will appear more inconsistent, the more specific are the response-equivalence classes used. Skinner (1938) obtained considerable response consistency and predictability using lever depression as a response-equivalence class. If, for example, the pressure applied to the lever had been used to divide lever presses into several smaller response-equivalence classes, then it is likely that most of this predictability would have vanished. Since response-equivalence classes are theoretically defined, apparent behavioural inconsistencies may be replaced by predictability when there is some theoretical understanding of the most appropriate response categories.

(4) Mischel (1973*a*) argued that traits are constructed from global overgeneralizations based on behaviour. He has not, apparently, considered the possibility that hereditary factors might be of importance. This is especially puzzling in view of the fact that the evidence from twin studies consistently indicates the substantial part played by heredity in the determination of personality. Shields (1962) carried out one of the most thorough investigations, and his study had the advantage of including monozygotic twins brought up apart. He used a fore-runner of the Maudsley Personality Inventory, and, for the extraversion scale, obtained intra-pair correlations of +0.61 for monozygotic twins reared apart, +0.42 for monozygotic twins reared together, and of -0.17 for dizygotic twins reared together. There was a similar pattern for neuroticism, with the correlations being +0.53 for monozygotic twins reared apart, +0.38 for monozygotic twins reared together, and +0.11 for dizygotic twins reared together. Although the low correlations for dizygotic twins and the greater correlation for monozygotic twins reared apart than together are somewhat problematical, the overall pattern of results is clearly indicative of some hereditary determination of personality traits. Jinks & Fulker (1970) reanalysed the data of Shields by the biometrical method of analysis, and obtained heritability estimates of 54 per cent for neuroticism and of 67 per cent for extraversion.

The experimental evidence from all the relevant twin studies was reviewed by Shields (1973), who concluded that nearly all the studies showed evidence of a significant hereditary component in extraversion, and many studies showed the same with respect to neuroticism or anxiety. Other reviews of the literature are available in H. J. Eysenck (1976*a*) and Nichols (1978).

In sum, it appears that Mischel has ignored a crucially important determinant of individual differences in personality, thus severely reducing the persuasiveness of his account of the origins of traits. A further important point is that, given the existence of a significant involvement of heredity in personality differences, any adequate theory of personality must take account of hereditary factors. It is not obvious how this could be done within the context of social learning theory (Mischel, 1973*a*). On the other hand, trait-state theories have typically emphasized the point that personality traits involve some hereditary component. Indeed, a critical issue in contemporary personality theory is (or should be) the role played by heredity. Since the evidence indicates that hereditary factors are important in explaining individual differences in personality, and since the trait-state approach is almost the only major theory of personality that acknowledges that fact and

incorporates hereditary factors by means of the trait concept, it is incumbent upon theorists of different persuasions to address themselves to this issue.

(5) Mischel (1968) pointed out that cross-situational behavioural measures rarely produce correlations in excess of +0.30, or 9 per cent of the variance, and thus it appears that behavioural inconsistency is the rule rather than the exception. While such relatively low reliability coefficients would undoubtedly embarrass a simple trait theory, there are two additional pertinent considerations. Firstly, in work discussed in more detail subsequently, Sarason *et al.* (1975) found in a review of almost 140 analyses of variance that personality accounted for approximately 9 per cent of the variance on average, and the situation for 10 per cent. If one adopts very stringent criteria for the minimal percentage of the variance that a factor must account for in order to warrant further consideration, then there is the danger that researchers will discover that no factors at all are sufficiently important to consider!

Secondly, Mischel (1968, 1973*a*) has recognized that the task of predicting behavioural responses within a trait-state theory can proceed on the basis of 'moderator variables' (Wallach, 1962). The basic notion is that the influence of any particular trait on behaviour will usually be indirect, being affected or 'moderated' by a number of other traits, mediating variables and situational factors. Mischel has criticized this approach, arguing that the more moderators that are required to qualify a trait, the more a trait-based formulation resembles a relatively specific description of a behaviour-situation unit. While it is true that trait-state conceptualizations have become increasingly complex over the last few years, it could very well be argued in view of the complexity of human functioning that this is a necessary, and indeed inevitable, development.

Evidence of some cross-situational specificity of behaviour can only be taken as highly damaging for state-trait theories that assume a direct one-to-one correspondence between internal traits and behavioural indices. Since most contemporary state-trait theories postulate the existence of moderator variables and thus claim only an indirect but theoretically predictable relationship between traits and behavioural responses, Mischel's evidence loses much of its apparent force.

(6) Although it is desirable, as Mischel has emphasized, that those factors emphasized by a theoretical position should account for a sizeable proportion of behavioural variation, it is also the case that there are various other criteria by which theories can, and should, be evaluated. One such criterion is a theory's range of applicability. In terms of that criterion, state-trait theories have often been outstandingly successful. For example, as H. J. Eysenck (1971, 1976*b*) has shown, the personality dimension of introversion-extraversion has been found to be related to performance in a theoretically predictable way in the following, and other respects: sensory threshold; pain threshold; time estimation; sensory deprivation; perceptual defence; vigilance; critical flicker fusion; sleep-wakefulness patterns; visual constancy; figural after-effects; visual masking; rest pauses in tapping; speech patterns; conditioning; reminiscence; and expressive behaviour. Mischel's criticisms suffer from the disadvantage of evaluating the state-trait approach from a rather limited perspective.

(7) Mischel (1968) discussed a variety of research findings that appeared to demonstrate highly specific situational effects on behaviour. This apparent inconsistency of behaviour contrasts with a persistent tendency among most people to regard others as possessing stable and typical behavioural patterns.

Mischel (1968) attempted to explain this apparent paradox by arguing that people tend to proceed far beyond the actual observation of some consistency in behaviour to the attribution of greater perceived consistencies which they construct. However, an alternative viewpoint is possible, starting with the fact that theoretical analyses and experimental paradigms have tended to incorporate the assumption that interrelationships between the

individual and the situation are unidirectional, i.e. the situation affects the person. Indeed, the typical laboratory experiment involves the manipulation of some experimenter-determined aspects of the situation (the independent variable) in order to observe the behavioural consequences. Since the subject does not control the situations to which he or she is exposed, only information about situational influences on behaviour can be obtained.

Wachtel (1973) has referred to this model of research, with the behaviour of the experimenter intended to occur independently of the subject's activities, as the model of the 'implacable experimenter'. Since it is, in fact, indisputable that there are bidirectional influences of the situation on the person and of the person on the situation, it may well be that experimental research systematically underestimates the consistency of behaviour that actually occurs under naturalistic conditions.

(8) Mischel and his critics even seem to regard the figure of 0.30 as an average measure of consistency as meaningful in some way; it is difficult to see how this can be. Essentially Mischel is trying to prove a negative, i.e. conduct is not consistent. This clearly is not possible; even if all (n) attempts to discover consistency had been failures, the possibility that the ($n+1$) attempt may be successful is not ruled out. To average, as he has done, successful and unsuccessful attempts is meaningless; success clearly depends on having a theory essentially pointing in the right direction, choosing tests which are both reliable and valid, and applying them to an appropriate population under appropriate motivational circumstances. If even one of these (rare and unusual!) preconditions is missing, the failure of the experiment says nothing about consistency of conduct.

This point has to be seen against the typical way in which research into personality is conducted. Without wishing to caricature the modal paper in this field, it would seem that a multiphasic test of personality traits is administered to a population of sophomores, and that then all the separate scores of this test are correlated with some criterion. Usually little by way of hypothesis is stated, and even when a hypothesis is allegedly tested, it does not usually specify a particular trait. Granted that one of the many traits measured by the multiphasic test is relevant to the 'hypothesis', the others most likely are not; yet the correlations with the criterion are usually calculated for all. Given a test measuring 16 traits (like the 16 PF), and given that one trait is strongly related to the criterion, while the other 15 are not, averaging all the observed correlations must inevitably give a low mean r ; it is this meaningless mean or modal figure that enters into such average figures as those quoted by Mischel.

Two other possibilities must be considered. It is possible that none of the observed correlations with the criterion is high; this does not demonstrate lack of consistency of conduct, but merely indicates that the hypothesis tested (assuming there to have been one!) is erroneous. No further deduction can be made about consistency of conduct. Granted that in any novel field most hypotheses are likely to be in error, failures are to be anticipated; they should not be averaged out together with the successes. It has been shown that many different deductions can be made from well-established theories, such as that linking introversion with cortical arousal (H. J. Eysenck, 1976*b*; M. W. Eysenck, 1977); these cannot or should not be argued away by averaging them with the numerous failures of other theories.

Another important possibility is that the many traits measured by the multivariate test are correlated, and give rise to a much smaller number of 'superfactors'; thus the California Psychological Inventory (CPI), which has 18 scales, can be shown to measure essentially two major personality dimensions, i.e. neuroticism and extraversion (Nichols & Schnell, 1963). The question now arises of whether the specific variance of the 18 scales contributes anything to the measurement of various criteria over and above the common variance summarized in the N and E scores. Reynolds & Nichols (1977) have shown that

the answer is in the negative; the two superfactors contribute all the valid variance to the prediction of the criteria used in their study. We thus have identical degrees of consistency (test-criterion) regardless of whether we use 18 scales or two. In the usual way we would divide this consistency by 18, obtaining very small values, when we should really divide by two, thus obtaining much higher values for consistency! Thus do the customary methods of analysis, hallowed by time but unjustified by psychometric criteria, artificially lower quite drastically the apparent consistency of conduct. Such statistical artefacts must be guarded against. We conclude that Mischel's summary figure for consistency is strictly meaningless.

(9) Even if we could take the figure of 10 per cent test contribution to the variance as a measure of consistency seriously, it would still not follow that personality factors were relatively unimportant. Criteria of conduct are usually highly complex, and any particular personality trait by itself would not be expected to predict such complex behaviour perfectly. Let us assume that particular behaviours are determined by 10 independent traits of personality, each contributing 10 per cent to the variance; this would give us perfect predictive accuracy, even though each trait obeyed completely the Mischel limitations! It is not suggested that personality testing would ever achieve such high predictive accuracy, but some such argument of additivity (or even multiplicability) of traits has been demonstrated. High N shows quite different behaviour when accompanied by high or low E (H. J. Eysenck, 1967); adding high or low P (psychoticism) will again change the outcome (H. J. Eysenck & S. B. G. Eysenck, 1976). Combining the particular combination of P, E and N with high or low intelligence will again powerfully determine the outcome, and so forth. Mischel criticizes a simplistic and unrealistic position never seriously taken by any personality theorist; he fails to come to grips with the actual position taken by Cattell, Guilford or Eysenck. As in chemistry, study of individual substances is a beginning, but it is not enough; we must also study their interrelations and combinations. No doubt the laws of combination present special difficulties, but the complexity of human (and animal!) behaviour requires us to abandon simplistic hypotheses and look at laws of combination which may be curvilinear, multimodal, and complex in many ways determined by the interaction of the traits originally measured. Mischel is aiming at a man of straw; he does not criticize modern personality theory as it really is.

(10) Mention has already been made of the need to specify the population tested in order to test properly a particular hypothesis. Thus the Hartshorne & May studies of honesty and deceit, in using young children in whom learning and conditioning would not yet have succeeded in producing very consistent reactions to complex situations, loaded the dice in favour of 'inconsistency'; that they still observed a considerable amount of consistency is evidence for the strong determination of conduct by personality. As H. J. Eysenck (1970) has shown, later workers, using similar tests with older subjects, obtained much stronger evidence for consistency. Again, averaging over many different populations as Mischel does is essentially a meaningless procedure, and the result cannot be interpreted in the way Mischel attempts to.

(11) It is well known in psychometrics that correlations cannot be interpreted directly without some knowledge of the internal reliability of the scores correlated. Any attempt to estimate the relationship between two variables which relies on unreliable estimates of one or both may grossly underestimate the 'true' correlation, and attempts should always be made to correct for attenuation. This is practically never done in the studies quoted and averaged by Mischel, although the reliabilities of the variables in question are often known, and frequently fall short of what might be regarded as adequate. For this reason, among others, the average correlation of 0.30 as a meaningful estimate of the 'true' relationships in question must be regarded as underestimating these to an unknown but probably substantial extent. Cattell *et al.* (1970) give tables of predictive correlations for many

criteria; the small size of the reported reliabilities of the tests used suggests that the 'true' relationships would be several times as strong as those reported.

(12) This argument becomes even stronger when we consider that much of the work criticized by Mischel attempts to predict real-life behaviour. Such predictions can come to grief for several reasons, among which is the obvious one that there is in fact no consistency of conduct such as is implied in the theory that reasonably high correlations will be found; this is in practice the only one considered by Mischel. It is equally possible, as has already been mentioned, that the wrong theory has been tested, or that the wrong tests have been used, or that the tests used were unreliable. What is equally possible, and often demonstrably true, is that the criterion may be excessively faulty, i.e. either invalid or unreliable. Educational criteria are famous for their lack of reliability (Hartog & Rhodes, 1936); other real-life criteria in industry, psychiatry and elsewhere often share this fault. Unless we have reason to believe that our criteria are both reliable and valid, the failure of tests to predict these criteria adequately cannot be used as proof of the inconsistency of conduct.

States and traits as intervening variables

A question that is fundamental to any assessment of the state-trait approach to personality is that of deciding upon adequate criteria for evaluating the adequacy of state and trait constructs. One possible criterion was favoured by Mischel (1969). He regarded cross-situational correlations as being of prime relevance, and concluded as follows: '[There is] impressive evidence that on virtually all of our dispositional measures of personality substantial changes occur in the characteristics of the individual over time, and, even more dramatically, across seemingly similar settings cross-sectionally' (p. 1012).

Mischel here disregards not only, as already noted, the powerful genetic evidence for hereditary determination of conduct; he fails to pay attention to the many studies demonstrating impressive consistency of conduct over time, often starting in early babyhood (e.g. Burt, 1965; Thomas & Chess, 1977). There are indeed changes; what is surprising is that there is also so much consistency. It is easy to be blinded by either the degree of consistency or of inconsistency, and to conclude that personality is more or less important in describing and determining conduct than it really is; such temptations must be resisted. In certain areas there is more longitudinal consistency than in others; we must seek quantitative evidence, but we must resist the temptation to calculate meaningless averages over all conditions, thus lumping the consistent areas with the inconsistent.

Since both state and trait concepts clearly represent intervening variables, it is reasonable to consider whether there are other ways of justifying the postulation of intervening variables, over and above the use of cross-situational correlations. A classic analysis of this issue was presented by Miller (1959). He pointed out that if an experimenter only has information about the effect of a single independent variable on a single behavioural measure (e.g. situational stress on difficult-task performance), then it is simplest to represent the relationship as a direct one of stress on performance (i.e. stress affects performance), since this account involves only a single functional relationship. Under these circumstances, the introduction of an intervening variable (e.g. anxiety) would merely complicate matters by requiring one to refer to two functional relationships (i.e. stress causes anxiety; anxiety affects performance).

In terms of the number of functional relationships required, the break-even point occurs when there are two independent variables and two dependent variables, since four functional relationships are required whether an account incorporating or not incorporating intervening variables is preferred. If an experimenter has information about three independent variables and three dependent variables, however, then two potential

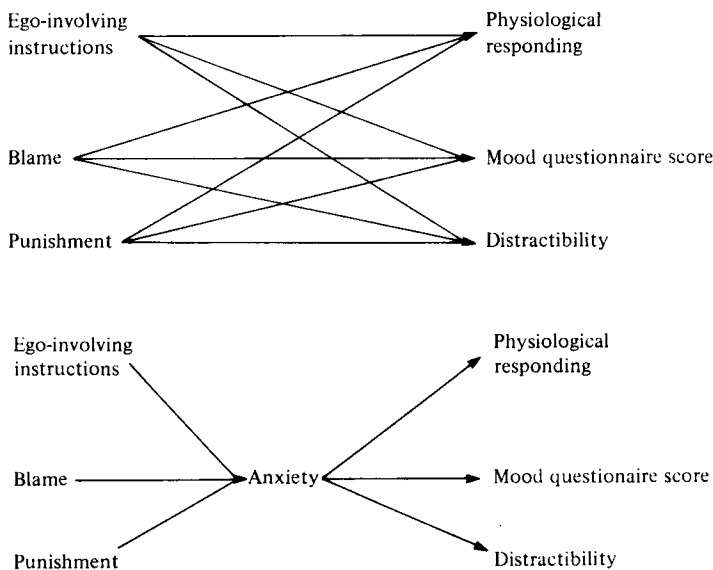


Figure 1. Two ways of describing interrelationships among independent and dependent variables: (a) without postulating an intervening variable; (b) with postulation of an intervening variable.

advantages can accrue from the use of a unifying intervening variable. Firstly, there is a real improvement in efficiency from nine functional relationships if no intervening variable is postulated down to six functional relationships if one is postulated (see Fig. 1). Secondly, the intervening-variable approach allows for experimental testing and possible disproof of the notion that a single intervening variable can account for the data. For example, if a given level of blame produces more distractibility than particular ego-involving instructions, then it should also produce greater anxiety on a mood questionnaire and stronger physiological responses.

This basic method of justifying state and trait constructs has been employed successfully with respect to several major personality dimensions, including extraversion and neuroticism (H. J. Eysenck, 1967) and state and trait anxiety (Spielberger, 1972). Use of this method would seem to provide a satisfactory way of refuting the oft-expressed criticism that trait concepts are inherently circular, an allegation forcibly put by Wiggins (1973): 'Perhaps the most objectionable feature of the trait construct... is the manner in which traits are construed as hypothetical entities which cause behaviour. The objection is... to hypothetical constructs that are animistic and circular, and that direct attention from lawful empirical relationships' (p. 366). As already indicated, once a trait construct is used to explain the diverse effects of several independent variables, as in Fig. 1, then that trait construct is clearly no longer circular or tautological. That this is the case can be seen from the fact that empirically testable predictions can be made.

Persons, situations, and their interaction

The fact that individual differences in personality often account for relatively modest percentages of behavioural variance has led several researchers to investigate whether situational factors might account for substantially larger percentages of the variance. In addition, there is an increasing awareness of the importance of interactions between the individual and the situation (e.g. Magnusson & Endler, 1977). Mischel (1973a) has argued that these interactions reflect idiosyncratic and theoretically unpredictable interrelationships.

The general expectations from the state-trait position are that situational factors will not prove to be substantially more consequential than person factors, and that many of the observed interactions between persons and situations are both replicable and theoretically predictable.

Several studies have been done in an attempt to evaluate the relative importance of the person and of the situation in determining observed behaviour. Bowers (1973) discussed 11 such studies that had calculated the percentages of the total variance accounted for by various factors. The mean percentage of the variance accounted for by persons was 11.27 per cent, compared with 10.17 per cent for situations, and 20.77 per cent for the interaction between persons and situations. Thus, the data suggested that situations and persons were of comparably modest consequence as behavioural determinants, whereas the interaction between the two factors was approximately twice as influential as either of the main effects.

A similar approach was adopted by Sarason *et al.* (1975), with the exception that they considered only studies reporting data on both personality and situational factors. The difference is that an interaction between persons and situations comprises a composite of all possible interactions between personality characteristics and situations for the particular situations of interest. Sarason *et al.* (1975) assessed the proportions of behavioural variance in each of 138 analyses of variance by means of the omega-squared statistic. On average, the situation accounted for 10.3 per cent of the variance, personality for 8.7 per cent, and the interaction between the situation and personality for 4.6 per cent. Thirty-five per cent of the situational main effects accounted for more than 10 per cent of the variance, whereas 29 per cent of the personality main effects and 11 per cent of the personality by situation interactions did so.

Before discussing the theoretical relevance of these data, it is necessary to point out their limitations. As Golding (1975) has noted, while the omega-squared ratios used in these studies do technically index the percentage total variation, they are inappropriate as measures of the general significance of the various factors. It would be possible, for example, for the reliability of a trait to be perfect across two situations, and yet still have individual differences account for only a small proportion of the variance. If the trait were running ability and the two situations were the 100 metres and the marathon, then clearly the dependent variable (running time in seconds) would be primarily affected by situational factors even if the rank order of performance were identical across situations. In other words, by suitable selection of situations and of personality dimensions, any pattern of results could be obtained.

In spite of these shortcomings, the data clearly suggest the inadequacy of any approach that fails to utilize both individual-difference and situational factors in the theoretical explanation of observed behaviour. A more challenging issue, however, concerns the proper interpretation of the interaction terms (i.e. persons by situations, and personality by situations). One extreme position on this issue was taken by Mischel (1973a): 'When interpreting the meaning of the data on Person \times Situation interactions and moderator variables, it has been tempting to treat the obtained interactions as if they had demonstrated that people behave consistently in predictable ways across a wide variety of situations. . . The available data on this topic now merely highlight the idiosyncratic organization of behaviour within individuals, and hence the uniqueness of stimulus equivalences for each person' (p. 258). We would argue, on the contrary, that replicated, theoretically predictable, interactions between individual-difference variables and situational factors have frequently been found, indicating that behaviour is by no means idiosyncratically organized within each individual.

One example of a consistently obtained interaction between a personality variable and a situation factor is discussed by M. W. Eysenck (1976). The Yerkes-Dodson Law stated

that task performance is interactively determined by arousal and task difficulty: the more difficult the task, the lower the optimal level of arousal. If one assumes that introverts are chronically more aroused than extraverts (H. J. Eysenck, 1967), then it is reasonable to predict that introverts will outperform extraverts on simple learning tasks, but that the reverse will occur on difficult tasks. This interaction has been obtained at least eight times (Siegman, 1957; Jensen, 1964; Shanmugam & Santhanam, 1964; McLaughlin & Eysenck, 1967; Howarth, 1969*a, b*; Bone, 1971; Allsopp & Eysenck, 1974).

Further examples of consistently replicated and theoretically predicted interactions between personality and situational factors are reviewed by Costello (1964). He cited six studies that had obtained a significant interaction between instructional stress (neutral versus ego-involving) and trait anxiety, with the more motivating instructions being detrimental to performance for high anxiety subjects and facilitatory for low and medium anxiety subjects. A similar interaction between a different stress factor (absence versus presence of blame) and trait anxiety was obtained in seven different studies discussed by Costello (1964).

Personality as an indispensable part of experimental and applied psychology

Our discussion thus far has been directed towards demonstrating that Mischel's criticisms of personality theory and application is either non-factual or even anti-factual. In this last section we shall argue that the inclusion of personality variables in empirical studies of psychological problems is not only permissible, but mandatory. The evidence is by now very strong that personality enters into predictions made in experimental, social, abnormal, industrial, educational and pharmacological psychology to such an extent, and in so predictable a manner, that it is possible to demonstrate the effects of neglecting it as a variable interacting with the 'main effects' of experimental manipulation (e.g. H. J. Eysenck, 1967, 1976*b*, 1978; Broadhurst, 1978). The effect is to reduce the 'main effects' portion of the total variance to a relatively small proportion, and to inflate the error variance to an unacceptable degree. When personality is included explicitly, and in line with theoretical prediction, in the design, then much of the so-called error variance is recognized as being truly main effects \times personality interaction variance, and hence useful in writing proper prediction equations. Many studies are cited in the references given above where the main effects part of the variance is completely absorbed in the error variance when personality is excluded from the experimental design, but where all the main effects appear in the interaction with personality. Such startling evidence cannot be omitted from any discussion of the consistency of personality; both consistent (trait) and inconsistent (state) measures of personality are implicated (M. W. Eysenck, 1977).

It is possible to go further than this and to suggest that many of the theoretical battles which have been fought in psychology owe their origin and their very intransigence to the fact that the organisms studied by the proponents of the divergent theories were genetically different. Thus Jones & Fennell (1965) and H. J. Eysenck (1967) have suggested that the great debate between the followers of Tolman on the one hand, and of Hull and Spence on the other, regarding the major laws of learning, may have been sparked off and sustained by the choice of rats of the 'emotional' strain of C. S. Hall by Tolman, and of rats of the 'non-emotional' strain by Hull and Spence! Emotional rats behave in a fashion that gives support to S-S theorists like Tolman, while non-emotional rats behave in a fashion that gives support to S-R theorists like Hull and Spence. Some evidence for this suggestion is provided in an experiment reported by Jones & Fennell (1965), and the suggestion that theoretical quarrels may be due to differential selection of (animal or human) subjects finds support in many other studies (e.g. H. J. Eysenck, 1976*b*).

It may be further suggested that many theoretical suggestions fail to be confirmed, or

that experimental findings arising from such theoretical suggestions fail to be replicated, because personality factors are not taken into account. Thus for instance Shigehisa *et al.* (1973) have shown that the failure of many investigators to obtain clear-cut evidence for or against the hypothesis that sensory thresholds can be lowered by increasing sensory input in modalities other than that being tested, was due entirely to their failure to separate their subjects into extraverts and introverts; extraverts do show this tendency, whereas introverts show the opposite tendency, thus demonstrating, as predicted, Pavlov's transmarginal inhibition. No replicability can be expected from studies that do not include consistent behaviour patterns (in this case, introversion-extraversion) in their design. Even more impressive than the mere empirical finding is the fact that the finding was predicted on the basis of personality theory; it seems difficult to argue that consistent behaviour patterns do not exist when a theory postulating such patterns can be used to generate predictions so amply confirmed.

What has been said here harks back of course to Cronbach's (1957) Presidential Address to the American Psychological Association concerning 'The two disciplines of scientific psychology'. This address received many plaudits, but was in effect disregarded by most experimental psychologists on the one side, and by most personality psychologists (to coin a term which embraces all those who use correlational methods in an attempt to study individual differences in traits, abilities, or attitudes). Mischel's critique of personality theory has encouraged this unfortunate break between the two disciplines of scientific psychology; it is our hope that this reply to his criticisms may go some way towards healing it. Experimental psychology (and social, educational, abnormal, industrial and any other kind of psychology) is unreasonably handicapped by disregarding personality factors in the widest sense; personality theorists are unreasonably handicapped by not paying attention to the major findings of experimental psychology, and gearing their conceptions to the concepts most widely used there.

Summary and conclusion

Of the various criticisms made of state-trait approaches to personality, the central one is that the actual inconsistency of behaviour contrasts with the prediction from the state-trait approach of behavioural consistency. However, since most state-trait theories argue that there is consistency at the intervening-variable level rather than at the behavioural level, this criticism is not damaging.

If state-trait theories cannot be satisfactorily evaluated by measures of cross-situational consistency, then the issue becomes one of proposing suitable criteria. Among those that merit consideration are the following:

- (1) Trait constructs should have a demonstrated hereditary component.
- (2) Trait and state constructs should account for the interrelationships among several independent and dependent variables.
- (3) Trait and state constructs should have a wide range of applicability.
- (4) Interactions between personality and situational factors should be theoretically predicted and replicable.

In essence, many state-trait theories emerge with credit when evaluated against the set of criteria discussed in this article. The problem is that Mischel and others have evaluated state-trait approaches from a limited and sometimes irrelevant perspective.

References

- | | |
|---|--|
| <p>ALLSOPP, J. E. & EYSENCK, H. J. (1974). Personality as a determinant of paired-associates learning. <i>Perceptual and Motor Skills</i>, 39, 315-324</p> | <p>BLOCK, J. (1977). Advancing the psychology of personality: Paradigmatic shift or improving the quality of research? In D. Magnusson & N. S.</p> |
|---|--|

- Endler (eds), *Personality at the Crossroads: Current Issues in Interactional Psychology*. Hillsdale, N.J.: Erlbaum.
- BONE, R. N. (1971). Interference, extraversion and paired-associate learning. *British Journal of Social and Clinical Psychology*, **10**, 284–285.
- BOWERS, K. S. (1973). Situationism in psychology: An analysis and a critique. *Psychological Review*, **80**, 307–336.
- BROADHURST, P. L. (1978). *Drugs and the Inheritance of Behaviour*. London: Plenum.
- BURT, C. (1965). Factorial studies of personality and their bearing on the work of the teacher. *British Journal of Educational Psychology*, **35**, 368–378.
- CATTELL, R. B., EBER, H. W. & TATSUOKA, M. M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire (16 PF)*. Champaign, Ill.: Institute for Personality and Ability Testing.
- COSTELLO, C. G. (1964). Ego-involvement, success and failure: A review of the literature. In H. J. Eysenck (ed.), *Experiments in Motivation*. Oxford: Pergamon.
- CRONBACH, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, **12**, 671–684.
- EPSTEIN, S. (1977). Traits are alive and well. In D. Magnusson & N. S. Endler (eds), *Personality at the Crossroads: Current Issues in Interactional Psychology*. Hillsdale, N.J.: Erlbaum.
- EYSENCK, H. J. (1967). *The Biological Basis of Personality*. Springfield, Ill.: C. C. Thomas.
- EYSENCK, H. J. (1970). *The Structure of Human Personality*. London: Methuen.
- EYSENCK, H. J. (1971). *Readings in Extraversion-Introversion*. Vol. 3: *Bearings on Basic Psychological Processes*. London: Staples.
- EYSENCK, H. J. (1976a). Genetic factors in personality development. In A. R. Kaplan (ed.), *Human Behavior Genetics*. Springfield, Ill.: C. C. Thomas.
- EYSENCK, H. J. (1978). The development of personality and its relation to learning. In S. Murray-Smith (ed.), *Melbourne Studies in Education*. Melbourne: University Press.
- EYSENCK, H. J. & EYSENCK, S. B. G. (1976). *Psychoticism as a Dimension of Personality*. London: Hodder & Stoughton.
- EYSENCK, M. W. (1976). Extraversion, verbal learning, and memory. *Psychological Bulletin*, **83**, 75–90.
- EYSENCK, M. W. (1977). *Human Memory: Theory, Research and Individual Differences*. Oxford: Pergamon.
- GOLDING, S. L. (1975). Flies in the ointment: Methodological problems in the analysis of the percentage due to persons and situations. *Psychological Bulletin*, **82**, 278–288.
- HARTOG, P. & RHODES, E. C. (1936). *An Examination of Examinations*. London: Macmillan.
- HARTSHORNE, H. & MAY, M. A. (1928). *Studies in Deceit*. New York: Macmillan.
- HARTSHORNE, H. & MAY, M. A. (1929). *Studies in Service and Self-Control*. New York: Macmillan.
- HARTSHORNE, H. & SHUTTLEWORTH, F. K. (1930). *Studies in the Organization of Character*. New York: Macmillan.
- HOWARTH, E. (1969a). Personality differences in serial learning under distraction. *Perceptual and Motor Skills*, **28**, 379–382.
- HOWARTH, E. (1969b). Extraversion and increased interference in paired-associate learning. *Perceptual and Motor Skills*, **29**, 403–406.
- JENSEN, A. (1964). *Individual Differences in Learning: Interference Factor*. Cooperative Research Project No. 1867, Office of Education, US Department of Health, Education and Welfare.
- JINKS, J. L. & FULKER, D. W. (1970). A comparison of the biometrical genetical, MAVA and classical approaches to the analysis of human behaviour. *Psychological Bulletin*, **73**, 311–349.
- JONES, M. B. & FENNELL, R. S. (1965). Runway performance by two strains of rats. *Quarterly Journal of Florida Academy of Science*, **28**, 289–296.
- MCLAUGHLIN, R. J. & EYSENCK, H. J. (1967). Extraversion, neuroticism, and paired-associates learning. *Journal of Experimental Research and Personality*, **2**, 128–132.
- MAGNUSSON, D. & ENDLER, N. S. (1977). Interaction psychology: Current issues and future prospects. In D. Magnusson & N. S. Endler (eds), *Personality at the Crossroads: Current Issues in Interactional Psychology*. Hillsdale, N.J.: Erlbaum.
- MILLER, N. E. (1959). Liberalization of basic S–R concepts: Extensions to conflict behaviour, motivation, and social learning. In S. Koch (ed.), *Psychology: A Study of a Science*, vol. 2. London: McGraw-Hill.
- MISCHEL, W. (1968). *Personality and Assessment*. London: Wiley.
- MISCHEL, W. (1969). Continuity and change in personality. *American Psychologist*, **24**, 1012–1018.
- MISCHEL, W. (1973a). Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, **80**, 252–283.
- MISCHEL, W. (1973b). On the empirical dilemmas of psychodynamic approaches: Issues and alternatives. *Journal of Abnormal Psychology*, **82**, 335–344.
- MISCHEL, W. (1977). The interaction of person and situation. In D. Magnusson & N. S. Endler (eds), *Personality at the Crossroads: Current Issues in Interactional Psychology*. Hillsdale, N.J.: Erlbaum.
- NICHOLS, R. C. (1978). Twin studies of ability, personality and interests. *Homo*, **29**, 158–173.
- NICHOLS, R. C. & SCHNELL, R. R. (1963). Factor scales for the California Psychological Inventory. *Journal of Consulting Psychology*, **27**, 228–235.
- REYNOLDS, C. H. & NICHOLS, R. C. (1977). Factor scales of the CPI: Do they capture the valid variance? *Educational and Psychological Measurement*, **37**, 907–915.
- SARASON, I. G., SMITH, R. E. & DIENER, E. (1975). Personality research: Components of variance attributable to the person and the situation. *Journal of Personality and Social Psychology*, **32**, 199–204.
- SHANMUGAN, T. E. & SANTHANAM, M. C. (1964). Personality differences in serial learning when interference is presented at the marginal visual level. *Journal of the Indian Academy of Applied Psychology*, **1**, 25–28.
- SHIELDS, J. (1962). *Monozygotic Twins Brought up Apart and Brought up Together*. Oxford: Oxford University Press.
- SHIELDS, J. (1973). Heredity and psychological abnormality. In H. J. Eysenck (ed.), *Handbook of Abnormal Psychology*, 2nd ed. London: Pitman.

- SHIGEHISA, P. M., SHIGEHISA, T. & SYMONS, J. R. (1973). Effects of intensity of auditory stimulation on photopic visual sensitivity in relation to personality. *Japanese Psychological Research*, **15**, 164–172.
- SIEGMAN, A. W. (1957). Some relationships of anxiety and introversion–extraversion to serial learning. Unpublished PhD thesis, Ann Arbor.
- SKINNER, B. F. (1938). *The Behavior of Organisms*. New York: Appleton-Century-Crofts.
- SPIELBERGER, C. D. (1972). Anxiety as an emotional state. In C. D. Spielberger (ed.), *Anxiety: Current Trends in Theory and Research*, vol. 1.
- THOMAS, A. & CHESS, S. (1977). *Temperament and Development*. New York: Brunner/Mazel.
- THORNDIKE, E. L. (1903). *Educational Psychology*. New York: Lemcke & Brechner.
- VERNON, P. E. (1964). *Personality Assessment*. London: Methuen.
- WACHTEL, P. L. (1973). Psychodynamics, behaviour therapy and the implacable experimenter: An inquiry into the consistency of personality. *Journal of Abnormal Psychology*, **82**, 324–334.
- WALLACH, M. A. (1962). Active–analytical vs. passive–global cognitive functioning. In S. Messick & J. Ross (eds), *Measurement in Personality and Cognition*. New York: Wiley.
- WIGGINS, J. S. (1973). *Personality and Prediction: Principles of Personality Assessment*. Reading, Mass.: Addison-Wesley.

Received 22 January 1979

Requests for reprints should be addressed to Dr Michael W. Eysenck, Department of Psychology, Birkbeck College, University of London, Malet Street, London WC1E 7HX.
Professor H. J. Eysenck is at the Institute of Psychiatry, University of London.

Editorial note: Professor Mischel has declined an offer of space for reply, stating that his position is sufficiently well documented to make this unnecessary. It is worth noting that the best recent summary of his views – which he feels are misrepresented by Eysenck & Eysenck – is to be found in an article ‘On the interface of cognition and personality: Beyond the person–situation debate’, *American Psychologist* (1979), **34**, 740–754.