Part I

Personality traits and intelligence

1 Personality and intelligence: psychometric and experimental approaches

H. J. Eysenck

Introduction

A discussion of the relationship between two concepts requires some definition, however preliminary, of these concepts. Intelligence, as I have argued (Eysenck, 1988), has three major meanings in contemporary writings: biological intelligence, psychometric intelligence, and social intelligence. Figure 1.1 illustrates these 3 meanings. Biological intelligence is concerned with the physiological, neurological, and anatomical bases of intelligence; the existence of such bases is mandatory given the strong genetic determination of individual differences in this field (Eysenck, 1979, 1985).

Psychometric intelligence is defined in terms of the IQ; it constitutes Spearman's (1927) *g* factor (general intelligence), plus the various primary factors isolated since (Eysenck, 1979). It is determined largely by biological intelligence, but is also affected by environmental and social factors, in the ratio of 7:3, roughly speaking. Finally, social or practical intelligence (Sternberg, 1985; Sternberg & Wagner, 1986) is essentially concerned with the application of IQ to success in life; this depends largely on IQ but also on a variety of other factors as shown in Figure 1.1 (Sarton, 1969).

In this chapter I have concentrated on the definition of intelligence in terms of IQ, although scientifically I would argue that biological intelligence is more securely based (Eysenck, 1982). However, too little work has been done on the relation between biological intelligence and personality to make such an endeavor feasible. Practical intelligence is by definition almost bound together intrinsically with personality, insofar as personality determines very largely the use a person makes of his intellectual gifts (Eysenck, 1988). Psychiatric disorders, alcoholism, impulsivity, addiction, and promiscuity can fatally impair a person's ability to use his IQ to the best advantage, and will hence lead to impairment of "practical intelligence." This use of the term "intelligence" is too broad to be scientifically useful, and will hence not be used in this chapter. For readers interested in this concept, a recent paper by Miller, Omens, and Delwadia (1991) contains ample material.

Personality will here be used in terms of a hierarchical trait model, that is, a model based on primary, first-order factors, correlating to make up higher order concepts

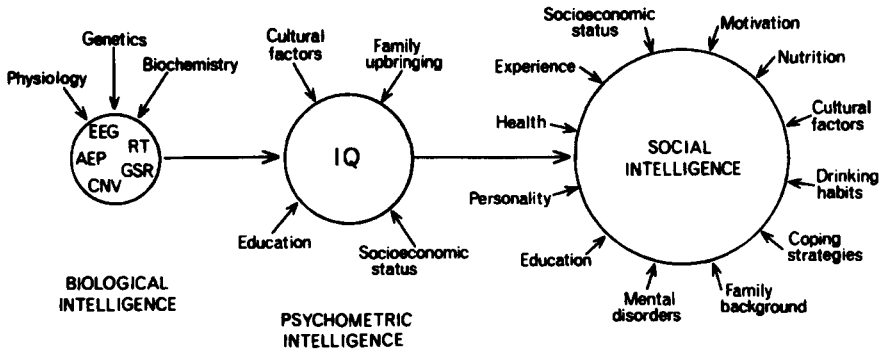

Figure 1.1. Different meanings of "intelligence." (From Eysenck, 1988.)

like extraversion (E), neuroticism (N) or anxiety, and psychoticism (P), as opposed to super-ego functioning (Eysenck & Eysenck, 1985). As regards the *relations* between these two concepts, I will emphasize the contrast between purely psychometric relations, a theoretical and purely heuristic, and a theoretical–experimental one, that is, based on experiments conducted to test a specific theory. I hope to show that little of importance has been found by authors using the first approach (unfortunately much the more numerous group), while potentially important findings have been made by authors using the second approach. The reasons why the theoretically and practically less promising approach has been so much favored is difficult to discern, unless it be that it is much easier, demands only routine collection of data, and can be done *ambulando.*

Intelligence and psychopathology

It is usually assumed, with good reason, that intelligence, as defined by Spearman's *g*, is not correlated with personality. A good deal of work has been done, for instance, on the possible relation between the WAIS, as a good measure of intellectual ability, and the MMPI, as a measure of personality. Gaines and Morris (1978) found that " . . . there are no lawful relationships between WAIS subtests and MMPI clinical scales" (p. 400). Similarly, Bloom and Entin (1975) concluded that "apparently no lawful relationships exist between WAIS and MMPI scales, and further investigation seems unwarranted" (p. 698). Again, Lacks and Keefe (1970) report lack of relationship, and conclude that "further pursuit of WAIS–MMPI relationships is apt to be fruitless" (p. 430).

On the other hand, Holland and Watson (1980), Turner and Horn (1976), and Watson, Davis, and McDermott (1976) do claim to have found relations between the two tests, as do Berg, Ingersoll, and Terry (1985). The last-named may serve to illustrate the possible degree of relationship that may be found. Using 197 psychiatric

inpatients, they administered MMPI and WAIS tests, and subjected valid raw subscale scores to bivariate, multivariate, and canonical correlational analyses. Twenty-four percent of the bivariate correlations and 50% of the multiple correlations were significant, and a canonical *R* of .609 was found between the WAIS subtests and the MMPI clinical scales, and a canonical *R* of .394 between the WAIS subtests and the MMPI validity scales. Tables given by the authors show the bivariate and multivariate correlations between MMPI and WAIS subtest scores.

How can we interpret the observed relations? The authors suggest that their canonical analysis has characteristics of a cognitive factor, for example, memory, for the clinical scales; for the validity scales, they suggest an intellectual factor. Holland and Watson (1980) had suggested an intellective factor to account for the clinical scales and their relation with the WAIS also, but of course all these interpretations are speculative. Perhaps all that these studies show is that when there are large differences in psychopathology, the degree (and possibly the kind) of pathology may depress WAIS scores selectively. For normal samples there should be little correlation, and that is what is usually found. Making avoidable errors seems correlated with high 428 and 987 MMPI profiles and may be the mediating factor between pathology and IQ (Fracchia, Fiorentiur, Sheppard, and Merlis, 1970).

Is there a relation between kind of psychopathology and intelligence? The MMPI is hardly an instrument to use when looking at differential pathology, due to its psychometric deficits, but Wechsler has from the beginning emphasized, on the basis of his clinical intuition, that certain patterns of WAIS profiles are in fact diagnostic of different pathologies (1940). Matarazzo (1972) has surveyed the literature; so have Rabin (1965) and Franks (1970). What did they find?

Matarazzo (1970) says, "Hundreds upon hundreds of studies on the use of profile, pattern, or scatter analysis with the Wechsler scales . . . fail to produce reliable evidence that such research would be fruitful." (p. 430). There would be little point in giving long lists of all these researchers with their negative or controversial findings, but it may be useful to consider why this obviously promising idea failed. The first reason, of course, is a psychometric one that should have been apparent from the beginning. The variables considered, whether they be profiles, patterns, or scatter analyses, are essentially measures of differences between subtests, and as such are subject to the damaging fact that differences between tests have much lower reliabilities than do the tests themselves. These lower reliabilities would make any meaningful findings very unlikely, and Meehl and Rosen (1955) and Cronbach and Gleser (1965) have presented the relevant statistical arguments in some detail; it is unfortunate that researchers in this area have neglected such considerations that, for all practical purposes, foredoom these empirical approaches to failure.

The second and equally important factor is the unreliability of the criterion, that is, the type of psychopathology involved. Psychiatric diagnoses are notoriously unreliable, and hence prediction of such unreliable diagnoses, even with a perfect test, would still show only very low correlations. Given this unreliability, Matarazzo points out that "a correlated index such as the Wechsler profile, no matter how

promising, cannot produce anything but . . . confusing, or otherwise frustrating findings."

There is a third reason why one would not expect such studies trying to use patterns or profiles on the Wechsler subtest as a diagnostic aid to succeed. Most of the studies were based on a simple blunderbuss approach; in the hope of finding correlations you administer a large number of tests and randomly correlate them with whatever takes your fancy, in anticipation of finding correlations (and judiciously forgetting to correct the observed probability values for the number of comparisons made!). This is the major fault of the WAIS-MMPI correlations; with so many correlations and no *a priori* hypothesis, the data are simply not worth calculating.

Clinicians often object to such criticisms by saying that their research is based on clinical "insights" which serve the purpose of a hypothesis. It is difficult to give credence to such arguments because the term "insight" is ill-defined. Usually it means little more than a guess on the part of the clinician, based on a certain amount of experience perhaps, but certainly not suggesting that his "insights" are anything but guesses. There is no theoretical rationale, based on sound laboratory evidence, and in the absence of such evidence the claim to possess some deep "insights" must be taken with a grain of salt. Matarazzo discusses in great detail the relationship between the Wechsler subtests and the Gittinger Personality Assessment System, which he regards as a theoretical model based on clinical insights. I know of no evidence to suggest that these "insights" have any relation to reality, or lead to any replicable relationships with Wechsler subtests.

It is fair, in retrospect, to deplore the time and energy spent by so many (mainly clinical) psychologists in trying to achieve something that in the nature of things was impossible. Rather than use unreliable and probably invalid clinical criteria (diagnoses or MMPI scores), they might have been better advised to try to improve the methods of diagnosis, particularly, replacing the categorical system with a dimensional one that seems to fit the facts much better (Eysenck, 1970). Furthermore, they might with advantage have improved their acquaintance with statistical methods and conclusions, thus avoiding the collection and analysis of data foredoomed to be inconclusive. Finally, they might have tried to elaborate more general theories that might predict relationships between intelligence and personality. This chapter is now pretty well closed, although some courageous souls are still searching for the Holy Grail in this unlikely area.

A slightly more hopeful relationship is that, originally suggested by Wechsler, between psychopathic personality and having a higher PIQ than VIQ. As Wechsler (1958, p. 176) put it, "the most outstanding single feature of the sociopath's test profile is his systematic high score on the Performance as compared with the Verbal part of the Scale." Matarazzo (1972) listed about 30 studies that generally support this view, although he urges caution in individual diagnosis. Moffit and Silva (1987) are less optimistic about the usefulness of this index. Kandell et al. (1988) find that high IQ acts as a protective factor for subjects at high risk for antisocial behaviour, but VIQ is equally protective as high PIQ. This agrees with the more general finding that IQ

correlates negatively with criminality (Eysenck & Gudjonsson, 1989). What is clearly needed is more systematic, focussed research of an experimental rather than a purely psychometric kind.

Personality and the structure of intelligence

It is usually taken for granted that the structure of intelligence is defined in terms of factorial investigation of correlations between tests. Thus, many studies have been done of the Wechsler scales (Maxwell, 1960; Canavan, Dunn, and McMillan, 1986), extending to abnormal groups, such as patients with unilateral cerebral lesions (Warrington, James, and Maciejewlski, 1986), and in many different countries (e.g., Drago, Daum, and Canavan, 1991) with similar results. Based on this assumption, we now have in addition to *g* some two dozen "primary factors" of intelligence, derived from many different tests (Eysenck, 1979). Similarly, the analysis of matrices of correlations between items of personality questionnaires, or of personality scales themselves, has given rise to numerous factors (Eysenck, 1970). Are these procedures justified?

The problem that arises may perhaps be put as follows: when a factor analysis of personality inventory scales is carried out, a number of factors, such as extraversion, neuroticism, etc., usually result. Similarly, when a factor analysis of intelligence test scales is carried out, a number of factors, such as verbal ability, perceptual ability etc., usually result. These factors are independent, in the first case of intelligence, in the second case of neuroticism or extraversion, as long as we preserve the rule that we are only concerned with linear relations. But we may enquire whether similar factors and relations would emerge if we extracted personality factors from populations differing in intelligence level, or intelligence factors from populations differing in degree of neuroticism.

A study by Shure and Rogers (1963) has attempted to answer the first question. They administered the eighteen scales of the California Psychological Inventory (CPI) to three student groups differing without overlap in IQ level, and then intercorrelated and factor analyzed the resulting scores for the three groups separately. They found that while there was considerable overall similarity in the solution, the total factor variance associated with their neuroticism factor dropped by over 30% in going from the high-ability group to the low-ability group. (The sum of squared loadings is, respectively, 5.18, 4.64, and 3.48 for the three groups.) No such change was observed in their extraversion factor, the sum of squared loadings being 3.46, 3.76, and 3.17, respectively, for the three groups. While confirmation is, of course, essential before too much credence can be given to this finding, it would appear that factorial studies of personality may not give invariant results under change of ability level.

The other problem raised is perhaps even more important from the educational point of view: would factorial studies of abilities be invariant under change of personality composition of the groups under analysis? The only paper concerned

specifically with this problem is one published by Lienert (1963). His work is based on 1,003 school children with a mean age of between 15 and 16; three-fifths of the children were male. These children were administered thirteen intelligence tests of the Thurstone type, constituting the so-called Leistungspruefsystem of Horn (1962a). Also administered was a personality questionnaire, modeled after Eysenck's Maudsley Medical Questionnaire by Horn (1962b) that gives a measure of neuroticism and also contains a lie scale. Seventy-seven subjects were excluded from the analysis because they had not completed all the tests or because of unusual lie scale scores. Of the remaining subjects, 259 labile and 262 stable children were selected as constituting the 25% highest-scoring and lowest-scoring subjects, respectively, on the neuroticism scale. There were no differences between the groups in age but there were more girls in the labile group. However, Lienert was able to show in a preliminary factor analysis that sex had no effect on the factorial structure of the tests. A product–moment correlation of the summed standard scores on the thirteen tests with neuroticism gave a value of −0.16. While statistically significant because of the large numbers, this is for practical purposes equivalent to a finding of orthogonality between the two variables.

Separate matrices of intercorrelations were calculated for the labile and stable subjects, respectively, and split-half reliabilities were calculated for all the tests for the two groups. Reliabilities did not differ, but the average intercorrelation of the tests was slightly and significantly higher for the stable group (.33 as opposed to .27).

Next, Lienert carried out a multiple factor analysis following Thurstone's procedure. It was found that eight factors could be extracted from the stable group and only four from the labile group. Communalities were lower for the labile than for the stable group and specific factors were more important for the labile than for the stable group. After rotation, it was found that three factors could be interpreted for the labile and six for the stable group; the latter were said to be closer to Thurstone's primary factors, whereas the former were much more mixed. These figures suggest strongly that children high and low on neuroticism differ very significantly in the way their mental abilities are structured. This conclusion is so important that a thorough critical analysis of the study seems in order.

Such a study was undertaken by Eysenck and White (1964), whose analyses in essence supported Lienert. Cohen and Wittemann (1967) attempted to replicate Lienert's study, and administered 14 standardized intelligence tests to 2,000 eighth-graders. They used a different neuroticism test from that used by Lienert, and their sample was more homogeneous with respect to age. They failed to support Lienert's results, the factor structure of high-, medium-, and low-neuroticism groups corresponding to a high degree. There the matter rests at the moment. It is sad that so important and fundamental a question is left in such an unsatisfactory state. It is perhaps typical of psychology that a theoretically important question like this only attracts a few students, while utterly meaningless clinical studies such as those considered in our second section are repeated time and time again.

Intelligence and introversion–extraversion

In this section I shall try to show how theory can suggest relationships between intelligence and personality, in contrast to the blunderbuss approach criticized in the first and second sections. Before discussing specific hypotheses, let us consider a distinction that is very important but seldom made in this field. Students of learning and conditioning will be familiar with the distinction between learning and performance; we may have learned a specific response, but whether that response will actually be made (performance) depends on many additional variables. Low problem solving in an IQ test is a measure of performance; personality may influence performance rather than abstract intellect, with measurable effects on the IQ. An IQ test lasts for up to 1 hour or more, and considerations of fatigue, vigilance, arousal, etc., may very well play a part.

Speed of working is another important variable that is closely connected with extraversion (Eysenck, 1967). Jensen (1964) has reported a study on 50 university undergraduates demonstrating the relevance of this factor. The Raven Progressive Matrices were administered to subjects individually, without time limit; however, the total time taken by the subject in doing the test was recorded by the tester without the subject knowing it. Correlations of E (extraversion) and N (neuroticism) with total score (−.13 and +.15) were insignificant, but E correlated −.46 ($p < .01$) with total time spent on doing the test. Raven scores were completely uncorrelated with time spent on the test. These findings suggest that had the test been given under timed conditions, the more extraverted subjects would have had a distinct advantage. It is this type of deduction that seems well worth testing, derived as it is from well-established theories. (N showed a completely insignificant correlation with speed of working.)

I have suggested that introverts are characterized by higher cortical arousal than extraverts, and hence show greater vigilance and less inhibition in extended performance tasks (Eysenck, 1967; Eysenck & Eysenck, 1985). This suggests a crucial test: extraverts should show a decline in performance on IQ tests during a lengthy administration. The first attempt to test this hypothesis was carried out by D. Furneaux, in an unpublished study (Eysenck, 1957).

The test used consisted of sixteen easy letter-series problems, preceded by two problems that were not scored, and followed by an insoluble problem, also not scored. The problems are so easy that in the population of university students tested errors do not occur; consequently, the score used is the time taken over each problem. The problems are roughly equal in difficulty, as determined by prior research. Under these conditions we would expect inhibition to affect the speed of work, and we would expect it to do so differentially for extraverts and introverts. No more precise prediction can be made, as the test which is being analysed was preceded by other tests of intelligence, thus making the situation too complex to allow precise prediction.

Subjects were given the Guilford personality scales, and 12 markedly extraverted and 15 markedly introverted subjects chosen for a comparison of their scores. The

most reliable type of score was found to be the rate of work for a given item, divided by the average rate of work for the whole test. The difference between the groups was found to be statistically significant, thus lending support to both our points—inhibition affects the rate of work, and it does so differentially for extraverts and introverts. Furneaux took one further step. Using a new sample of 130 students, he plotted each subject's scores and gave him a new score (pattern score) according to the degree to which his pattern of scores approached the extraverted or the introverted pattern. These pattern scores would be expected to correlate with extraversion–introversion scores on the questionnaire, and indeed a very significant correlation of .35 is reported by Furneaux, thus cross-validating the differential patterns found (Eysenck, 1957, pp. 132–133).

In a replication of the Furneaux experiment, Eysenck (1959) predicted that in the process of solving the 60 problems of the Morrisby Compound Series Test (Morrisby, 1955), a nonverbal intelligence test, extraverts would show greater reactive inhibition, and consequently a falling off in performance during the last quarter of the test as compared with the first three quarters. From 137 adult male and female neurotics, who were given the Maudsley Personality Inventory, were then chosen an introverted group (*E* score of 16 or below) and an extraverted group (*E* score of 30 or above). Nineteen extraverts and 28 introverts were available for testing. They were administered the test individually, without time limit, and each item was separately timed.

There were no differences in the total number of items correctly solved, or in the speed with which all items were finished. There was, however, a significant difference in the speed with which correct solutions were produced. On the first 45 problems, introverts were insignificantly slower than extraverts; on the last 15 problems, extraverts were significantly slower than introverts. When we turn to the speed with which items were abandoned unsolved, we find that there were no significant differences on the first 45 problems, but that on the last 15 problems extraverts gave up significantly more quickly. (A one-tail test was used for this comparison because the outcome had been predicted.) It is concluded that extraverts show greater work decrement on an intelligence test by taking longer to obtain correct solutions toward the end of the test, as compared with introverts, and by giving up more easily toward the end.

These studies show that general laws, such as those linking vigilance with introversion, extend to performance on IQ tests; problem solving behaves just like other types of performance, cognitive and noncognitive (Eysenck, 1967). It is interesting that in such attempts to apply general psychological principles to the specific problems of intelligence, personality interactions have been so rare; this would seem to be an interesting and rewarding field.

A rather different theory-driven approach has been pioneered by Robinson (1985). As he says, "Previously reported findings indicate that variation of EEG evoked-potential parameters is strongly related to both personality and intelligence differences" (Robinson, 1982a). These data and the associated theory imply that personality should relate to intelligence-test performance. Results are described in this