What IQ Tests Test

Ken Richardson
OPEN UNIVERSITY

ABSTRACT. As we approach the centenary of the first practical intelligence test, there is still little scientific agreement about how human intelligence should be described, whether IQ tests actually measure it, and if they don't, what they actually do measure. The controversies and debates that result are well known. This paper brings together results and theory rarely considered (at least in conjunction with one another) in the IQ literature. It suggests that all of the population variance in IQ scores can be described in terms of a nexus of sociocognitive-affective factors that differentially prepares individuals for the cognitive, affective and performance demands of the test—in effect that the test is a measure of social class background, and not one of the ability for complex cognition as such. The rest of the paper shows how such factors can explain the correlational evidence usually thought to validate IQ tests, including associations with educational attainments, occupational performance and elementary cognitive tasks, as well as the intercorrelations among tests themselves.

KEY WORDS: affective, cognition, intelligence, performance

Tests of 'intelligence', providing an 'intelligence quotient', or IQ, of individuals, as an estimate of their relative cognitive abilities, have formed a prominent aspect of the psychological sciences for nearly a century. To this day they are used very widely for purposes of educational and occupational prognosis, and for clinical diagnosis, as well as psychological research and theorizing. Coupled with claims that differences in scores reflect innate differences in cognitive ability, IQ testing has formed the basis for heated controversies about the distribution of ability potential, and thus for educational and social policy (Devlin, Feinberg, Resnick, & Roeder, 1997; Herrnstein & Murray, 1994). Assumptions about what IQ measures have recently encouraged the search for 'genes for IQ' in an expensive and highly publicized research programme (Plomin, 1999). Yet all this is unusual, scientifically, because there is still little scientific agreement about the source of variance represented by IQ scores (Mackintosh, 1998). Thus the question 'What do IQ tests test?' would seem to be of continuing, if not increasing, scientific importance. In this paper, I try to bring to bear on this question the

THEORY & PSYCHOLOGY Copyright © 2002 SAGE Publications. VOL. 12(3): 283–314
[0959-3543(200206)12:3;283–314;024012]

results of recent research and theoretical perspectives usually neglected (at least in conjunction with each other) in the IQ literature.

Historical Puzzlement

The different facets of the problem reduce to three main questions: What is human intelligence? Does IQ measure it? And, if not, what does it measure? The great difficulties encountered by attempts to answer them are now well known, so brief illustrations will suffice. They first came to light in a symposium in 1921 in which the editors of the *Journal of Educational Psychology* asked 17 leading theorists to state what they considered 'intelligence' to be, and by what means it could best be measured by group tests. The diversity of answers received, and the absence of agreement among them, have been famous ever since. They led to the half-joking, half-exasperated claim that 'intelligence is what intelligence tests test' (Boring, 1923, p. 35).

This exercise was repeated more recently by Detterman and Sternberg (1986). They wrote to a group of theorists, asking them the same questions that were put to the experts in 1921. Sternberg and Berg (1986) analysed the results for frequencies of mentioned attributes. Of the 25 attributes of intelligence mentioned, only 3 were mentioned by 25 per cent or more of respondents (half of the respondents mentioned 'higher level components'; 25 per cent mentioned 'executive processes'; and 29 per cent mentioned 'that which is valued by culture'). Over a third of the attributes were mentioned by less than 10 per cent of respondents (only 8 per cent of the 1986 respondents mentioned 'ability to learn').

The problem has continued. For example, a task force set up by the American Psychological Association to report on 'knowns and unknowns in the intelligence debate' (the APA group; Neisser et al., 1996) seems to maintain that IQ is a valid measure of intelligence, but warns that '[i]n this contentious arena, our most useful role may be to remind our readers that many of the critical questions about intelligence are still unanswered' (p. 97). They point out that views of what IQ measures range from a unitary general factor, to complex hierarchical ability structures, made up, in some views, of dozens of separate abilities. Comprehensive accounts of theory and findings in the area (e.g. Mackintosh, 1998) still produce no firm answers to the questions above. And Jensen (1998), who has been one of the foremost supporters of IQ as a measure of 'intelligence', now warns that 'the word "intelligence" has proved to be undefinable . . . without a scientifically acceptable degree of consensus' (p. 45).

In the absence of scientific characterization of human intelligence, the reverse logic scorned by Boring—i.e. that intelligence is whatever the source of variance in IQ scores actually is—has persisted, and statements about what IQ measures remain largely intuitive. The dominant intuition is that of a pervasive biological factor, permeating 'all aspects of cognition' (Gottfredson, 1998), and varying in the population as a biometric or simple quantitative trait like physical strength, energy or power. This is the conception promulgated by Galton (1883) and Spearman (1927), who called this factor *g* (for 'general ability'). However, *g* seems to remain as inscrutable as 'vital forces' once were in biology: as Jensen (1998) notes, the conception remains entirely metaphoric. It is largely sustained, not by an understanding of the source of variance per se, but from secondary evidence linking IQ scores with variance in what are assumed to be other manifestations of intelligence. What I shall argue below is that the source of population variance in IQ scores is far less mysterious, and more tractable, than such metaphors suggest. Some indication of this is gained from an examination of how IQ tests are actually constructed.

How IQ Tests Are Constructed

Danziger (1990) has documented the circumstances under which IQ testing became popular in a nascent discipline in the early part of the 20th century. Burgeoning state education systems on both sides of the Atlantic, largely geared to demands of economic efficiency, created the need for simple test data useful for administrative expedience. Opportunistic responses to this new 'market', as Danziger explains, shaped and redefined much of psychology. The aims of identifying individuals as quantifiable resources, more or less worthy of future investment, cut across the theory-based accounts of child psychologists like Hall, Dewey, Baldwin and others in that they emphasized the passivity and outward performance of the child, with little exploration of mental processes and wider contexts of their development. Following the axiom that all variation could be reduced to scalar differences, individual differences became defined in terms of test performances. The notion of a unitary, biologically fixed, general ability served that purpose very well. It provided the justification for 'grading the entire population as though they were members of one school class' (Danziger, 1990, p. 109), reducing variation to stable constitutional factors within the individual, with little reference to psychological states in historical and social contexts.

As a result, IQ tests are unusual instruments in several respects. Ordinary scientific measurement of concealed inner states, whether of inorganic, organic or psychological systems, is usually based on a widely agreed characterization or clear theoretical model of the internal processes involved in that state: in particular, how its parameters are reliably manifested in some external index that can be monitored, and against which an instrument can be calibrated. In this way we expect breathalysers to be valid tests of blood alcohol levels; or white cell counts to be valid measures of pathogenic infection. In each of these some causal connection between changes in internal parameters and changes in the index (or instrument for measuring it) is demonstrated, understood and agreed. Of course, measurement in the exploratory phases in any natural process may be highly tentative. As Terman (1916) put it in the case of IQ testing:

> The best that can be done . . . is to make tentative assumptions as to the probable nature of intelligence, and then subject these assumptions to tests which will show their correctness or incorrectness . . . and thus gradually [lead us] to a conception of intelligence which will be meaningful and in harmony with all the ascertainable facts. (p. 44)

Unfortunately, in the case of IQ, this testing of assumptions has not been done, and little scientific understanding of the relevant parameters of the human intelligence system has been worked out. In its absence, the strategy has been to work backwards—to take what is presumed to be an external index (the 'criterion') of intelligence and then evoke and select, by trial and error, performances of individuals that display population variance that covaries with the criterion. The sum total of these evoked performances is then taken to be a measure of intelligence. Galton (1883) proposed social status as the criterion, but the performances he evoked, using tests of sensorimotor parameters like reaction times, perceptual accuracy, and so on, failed because population variance in them did not covary with the criterion.

The first successful attempt to implement this 'backwards' strategy for estimating intelligence was that of Alfred Binet in 1905. Binet's aim was to screen out children in Parisian schools who were 'backward' (for whatever reason) so that they could be given special treatment (as stipulated in new legislation). With his assistants, Binet spent long periods working in schools observing and analysing the kinds of knowledge and skills children were expected to learn and on which they were assessed (Miller, 1962). This 'fieldwork', along with intense discussions with teachers and other educators, convinced him that an adequate predictive instrument should combine multiples of school-type tasks involving more complex 'mental' constructs than those tried by Galton. So Binet devised short questions or problems on memory for specific facts; defining common words; verbal and arithmetical reasoning; counting; sentence construction; arranging sets of blocks in order of increasing size; and so on.

These efforts, as Miller (1962) notes, were pragmatic rather than scientific: items were accepted or discarded not on theoretical grounds, but purely on the extent to which individual performances on them agreed with the external criterion of school progress. The extent of this covariation was assessed by administering candidate items to a group of 'normal' children and a group of 'backward' children (both groups of various ages) as identified by teachers and doctors. The numbers of correct answers to the problems were then compared across groups. The differences were not great until Binet also added the criterion of age, resulting from his observation that backward children could often perform on the items as well as normal children, but at a later age. His first test incorporating this criterion consisted of 30 items arranged into bands corresponding to ‘expected’ performances of normal children in each age group from 3 to 11 years (see Zenderland, 1998). A ‘mental age’ could then be estimated for each child according to how far he or she progressed through the series. The procedure, then, identified children who were backward for their ages on items devised for their relevance to scholastic attainment.

This pragmatic procedure, using a ‘backwards’ strategy for arriving at a test that predicts school attainment, is the essence of the ‘Binet method’, and it has characterized IQ test construction ever since. Binet demurred about what exactly was being revealed by mental age scores—what the source of variance is variance in. Indeed, he insisted that he was not ‘measuring’ in any scientific sense at all: ‘Psychologists do not measure . . . we classify’, he said (quoted by Zenderland, 1998, p. 96). Specifically, he saw his test as a means for ‘classifying’ children along stages of mental development through which everyone passes.

Elsewhere, though, the tests were quickly accepted as genuine tests of ‘intelligence’ in the Galton/Spearman sense (Zenderland, 1998). Goddard (1911) translated Binet’s items for use in the USA. Terman (1916) devised 40 additional items to form the new Stanford–Binet test, which has since, over several revisions, become the ‘standard’ IQ test. The other most popular test, the Wechsler (1958) scales, was based broadly on the kinds of items that had been found to ‘work’ in previous tests (Anastasi, 1990). New types of items—especially non-verbal, or so-called ‘performance’ items—have been devised in the same pragmatic way, with little improvement in theory about cognitive processing to guide it (see further below). As Anastasi (1990) said about the Wechsler scales, ‘The weakest feature . . . has been their lack of theoretical grounding, which makes it hard to find a coherent basis for interpretation’ (p. 222). Instead, score variance has been attributed in the main to variance in the hypothetical/anonymous entity *g* (Jensen, 1998). In the rest of this article I try to show that this tacit theory is mistaken, and that there are more tractable sources of variance in IQ scores than that.

The Source(s) of IQ Variation

I shall argue that the basic source of variation in IQ test scores is not entirely (or even mainly) cognitive, and what is cognitive is not general or unitary. It arises from a nexus of sociocognitive-affective factors determining individuals’ relative preparedness for the demands of the IQ test. These factors include (a) the extent to which people of different social classes and cultures have acquired a specific form of intelligence (or forms of knowledge and reasoning); (b) related variation in 'academic orientation' and 'self-efficacy beliefs'; and (c) related variation in test anxiety, self-confidence, and so on, which affect performance in testing situations irrespective of actual ability.

Special Cognitive Preparedness

The IQ test is predicated on a simple biometric model of intelligence. Although essentially uncharacterized, the model assumes a domain-general computational function with population variance arising in simple parameters like 'capacity', 'speed' and 'efficiency' (Jensen, 1998). Likewise, problem-solving ability across individuals varies with those parameters irrespective of the problems' history and current context. However, alternative theorists, from Binet (see Zenderland, 1998) to contemporary developmentalists, have maintained the need for more 'psychological' models of cognition. These include recent 'ecological' (e.g. Wozniak & Fischer, 1993) and sociohistorical perspectives (e.g. Cole, Engeström, & Vasquez 1997; Richardson, 1998; Wertsch, del Río & Álvarez, 1995). In these perspectives, advanced cognitive systems evolved for coping with highly changeable, unpredictable circumstances, especially those that arise—or, indeed, are *created*—in social contexts (Donald, 1991). These are not dealt with by stereotyped computational mechanisms but by distilling the abstract informational structure (i.e. knowledge) peculiar to each problem domain. Such representations can then be used as 'psychological tools' to generate processes unique to the problem, but informed by that knowledge structure. Intelligent performances are thus based on an interaction between the 'structure' of the current problem and developed representations (i.e. knowledge), rather than stereotyped 'reasoning processes' that merely vary in speed or efficiency.

An important aspect of this perspective is that the complexity of structure of problems is vastly amplified in humans by being embedded in socially cooperative activities (Donald, 1991). In general, these are organized and regulated by *socially* evolved 'cultural tools'. Cultural tools include historically evolved patterns of co-action; the informal and institutionalized rules and procedures governing them; the shared conceptual representations underlying them; styles of speech and other forms of communication; administrative, management and accounting tools; specific hardware and technological tools; as well as ideologies, belief systems, social values, and so on (Vygotsky, 1988).

In sum, cultural tools prevail as the very means of human existence in any society. In sociohistorical theory, the forms of these cultural tools become internalized in individuals as the dominant 'psychological tools' (Vygotsky, 1988). Culture is thus more than mere clothing on human cognition, and forms of intelligence are more than something merely ‘valued’ by a culture, as both Gardner (1983) and Sternberg (e.g. 1999) suggest. Rather, culture is constitutive of its form and function—the ‘technologies of the intellect’, as Olson (1986) calls them. As Cole put it, ‘The structure of thought depends upon the structure of the dominant types of activity in different cultures’ (quoted by Luria, 1976, pp. xiv–xv). In addition, culture incorporates an already self-variegating, adaptable, cognitive system into a further system of variegation. In consequence, sociocognitive variation is both fundamentally different from, and vastly greater than, variation in a simple quantitative trait (and for very good evolutionary reasons). Generally speaking, this means that intelligence differentiates both ‘horizontally’ and ‘vertically’: horizontally, in variegating qualitatively, like languages or the products of the immune system; vertically, in the sense of degree of developed acquisition of specific cultural tools.

The IQ test collapses this rich and complex variegation in human cognition into a single scale as follows. In societies organized through hierarchical divisions of labour, different social classes will (by definition) utilize the different ‘cultural tools’ of society to different extents. Parents will thus vary in the degree to which they are cognitively enfranchised in the use of different cultural tools, so that their children will be prepared for acquiring them to varying degrees (and with different degrees of importance and emphasis—see below). Yet IQ tests, the items of which are designed by members of a rather narrow social class, will tend to test for the acquisition of a rather particular set of cultural tools: in effect, to test, or screen, for individuals’ psychological proximity to that set per se, regardless of intellectual complexity or superiority as such. I will try to illustrate this briefly with reference to some of the most typical IQ test items.

For example, most IQ tests, like the Stanford–Binet and the Weschler scales, include numerous questions like ‘What is the boiling point of water?’; ‘Who wrote Hamlet?’; ‘In what continent is Egypt?’; and so on. These most clearly demand little more than rote reproduction of factual knowledge most likely acquired, as valued pieces of information, from experience at home or by deliberate teaching in school. However, much research indicates how opportunities and pressures for acquiring such valued pieces of information—from books in the home to parents’ interests and educational level—are more likely to be found in middle-class than working-class or ethnic minority homes (for reviews see Mackintosh, 1998; Martinez, 2000). It has also been observed how mothers with more formal education tend to mimic teachers in asking their children more ‘known-answer’ questions (Greenfield, 1998). Such individuals will thus be much better prepared for an IQ test purely by virtue of their cultural background rather than cognitive ability as such. Yet such items are abundant in the most popular IQ tests and contribute much to the final score.

Other common IQ test items demand that culturally quite specific forms of language have been acquired. This is most obvious with items like verbal comprehension (sentence completion). But even 'verbal-mathematical' and 'performance' items are presented in heavily coded language forms (e.g. 'Why do we have clocks?'; 'This is a cut-up horse. Put the pieces together as quickly as you can'). Culturally specific ways of coding semantic relations constrain ways of interpreting and thinking about problems, which can, in turn, affect performance on them. As Olson (1986) notes, 'It is easy to show that sensitivity to the subtleties of language are crucial to some undertakings' (p. 341). He illustrates this with another prominent set of IQ test items consisting of word definitions: for example, 'What does Affliction mean?' These demand a:

> . . . highly metalinguistic form of analysis. . . . [T]he form of the definition is not conversational, as would be the case if one were merely seeking information. Rather, it requires a particular linguistic frame, an equative verb of being, is, and a syntactic formula, NP is NP, . . . [which] is not a natural linguistic form. (p. 350)

Take, as another example, simple verbal analogies that are used in some IQ tests because performances on them correlate with the external criterion. Mackintosh (1998) suggests that performance on analogical reasoning items like:

> black is to white as night is to (day)

illustrate Spearman's (1927) theory of intelligence as the 'eduction of relations and correlates' (the relation 'opposite' is educed and the answer deduced from the eduction of the correlate). We thus seemingly have to conclude that children who fail such a task cannot 'see' that day is opposite to night in the way that white is opposite to black. It could be, however, that some children simply have problems with the linguistic form 'is to'.

Numerous studies have shown how acquisition of these specialized psycholinguistic forms depends on what goes on in the home and local culture. Particular patterns of verbal interaction between parents and children that are more likely in middle-class homes are strongly related to IQ test performance (Hart & Risley, 1995, cf. Martinez, 2000). Sigel, Stinson and Kim (1993) found that the way that parents ask questions, and the dialogues they set up, influence how children 'learn to plan and guide their own behavior' and 'internalize the learning strategies and problem-solving approaches employed by parents' (p. 222). Other studies have shown how middle-class, much more than working-class, parents socialize their children into the use of written materials long before they start school, as 'an integral part of communication, recreation and livelihood . . . and the texture of daily life' (Rogoff, 1993, p. 87). This no doubt reflects the differential prominence of textual cultural tools in parents' own working lives. Even differences of detail in style of reading are later reflected in the children's abilities to analyse and comprehend textual material in school (Heath, 1982).

Although these culturally conditioned aspects of cognitive performance are sometimes acknowledged by those who favour IQ tests (e.g. Jensen, 1998), they respond by pointing out that parallel variance is found on scores on other, 'non-verbal', items. It is claimed that such items do not tap any culturally acquired knowledge or psycholinguistic forms, so the persistent variance refutes the idea that IQ is nothing more than degree of exposure to middle-class culture (Mackintosh, 1998).

Such an argument may be shallow, however, because it neglects any deeper analysis of the cognitions actually needed to solve such items. Take, for example, the Raven's Progressive Matrices (RPM) test (Figure 1). According to many authors (e.g. Jensen, 1980), the RPM is an almost 'pure' measure of 'intelligence' or *g*. Carpenter, Just and Shell (1990) put it close to the 'centre of gravity' of whatever it is IQ tests test. But, again, this is not based on an objective description of the cognitive processes required by, and exercised in solving, the items. Although it is true that the development of the RPM was based on Spearman's (1927) notion of intelligence as the 'eduction of relations and correlates', this does not necessarily specify complex cognition: the eduction of relations and correlates is exhibited in quite simple animals (Kesner & Olson, 1990). Indeed, the relations that appear in the RPM items are not, in fact, very complex (see below). Accordingly, as Carpenter et al. (1990) explained after examining Raven's personal notes, 'the description of the abilities that Raven intended to measure are primarily characteristics of the problems, not specifications of the requisite cognitive processes'. Judgements of complexity remain essentially impressionistic, just as Raven 'used his intuition and clinical experience to rank order the difficulty of the six problem types . . . without regard to any underlying processing theory' (p. 408).

Only a little further analysis suggests that the cognitive processing demanded by Raven's items simply reflects knowledge structures most common in one particular culture. Thus, many middle-class cultural tools are based on the manipulation of symbols (e.g. words, numbers) in two-dimensional array on paper. These include record sheets, tables with rows and columns of totals and subtotals, spreadsheets, timetables, and so on, as well as textual material. These nearly all require the reading of symbols from top left to bottom right, additions, subtractions and substitutions of numbers or other symbols across columns and down rows, and the deduction of new information from them. As the analyses of Carpenter et al. (1990) show, these are precisely the kinds of manipulations (or 'rules') built in to Raven's items (see Figure 1).

So although the symbols in a test like the RPM are experience-free, the rules governing their changes across the matrix are certainly not, and they are more likely to be already represented in the minds of children from

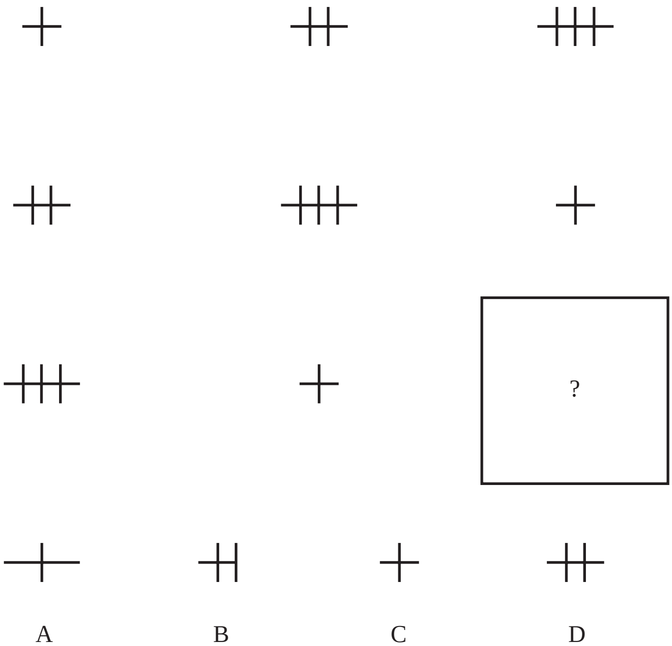

FIGURE 1. A matrix problem typical of ones used in the RPM (testees are required to select the ‘missing’ element from the options beneath).

middle-class homes, less so in others. Performance on the Raven’s test, in other words, is a question not of inducing ‘rules’ from meaningless symbols, in a totally abstract fashion, but of recruiting ones that are already rooted in the activites of some cultures rather than others. Like so many problems in life, including fields as diverse as chess, science and mathematics (e.g. Chi & Glaser, 1985), each item on the Raven’s test is a recognition problem (matching the covariation structure in a stimulus array to ones in background knowledge) before it is a reasoning problem. The latter is rendered easy when the former has been achieved. Similar arguments can be made about other so-called ‘culture-free’ items like analogies and classifications (Richardson & Webster, 1996).

Such cultural (background knowledge) facilitation of reasoning has now been demonstrated with respect to a large variety of cognitive tasks in which subjects can map the covariation relations in the tasks on to relations already represented in their background knowledge. These include the Wason selection task (Cheng & Holyoak, 1985); computerized ‘games’ governed by complex ‘rules’ (Ceci & Roazzi, 1994); analogical reasoning tasks (Goswami, 1996; Richardson & Webster, 1996); class-inclusion and scientific reasoning tasks (Carey, 1988; Franklin, 1997); categorization tasks (Keil, 1988); and modified Raven's matrices in which the covariation structure of the originals can be related to socially meaningful contexts (Richardson, 1991, 1996). In contrast, it has been shown how the degree of specialized knowledge required for competence in non-Western cultures has little association with performance on tests like the RPM (e.g. Sternberg, 1999). All these findings reveal a substantial gap between individuals' 'functional' performance on unfamiliar tests and their 'optimal' performance in familiar contexts in which they are well prepared (see Fischer, Bullock, Rotenberg, & Raya, 1993).

Such analyses also explain why the Raven's (and other non-verbal tests), often referred to as 'culture-free', 'culture-fair', 'culturally reduced' or tests of 'fluid intelligence', are, in fact, the most enculturated of all IQ tests (Keating & MacLean, 1987). Since all human cognition takes place through the medium of cultural/psychological tools, the very idea of a culture-free test is, as Cole (1999) notes, 'a contradiction in terms . . . by its very nature, IQ testing is culture bound' (p. 646). Individuals are simply more or less prepared for dealing with the cognitive and linguistic structures built in to the particular items.

All this is reinforced by other analyses showing that most IQ test items—even the most difficult items of the Raven's test—are not very complex cognitively. As Carpenter et al. (1990) explained, in their analysis of the Raven's items, 'Interestingly, the level of abstraction of even the most difficult rule . . . does not seem particularly great compared with the abstractions that are taught and acquired in various academic domains such as physics or political science' (p. 428). Everyday social interaction and practical problems seem to involve much greater cognitive complexity, and this might explain why there seems to be little if any little relationship between IQ scores and ability to solve complex sociocognitive problems in real life (see further below).

As mentioned above, it is likely that the cultural processes underlying this cognitive preparedness are quite subtle: learning, knowledge and reasoning ability having more to do with the structural aspects of cultural tools, and exposure to them, than general cognitive stimulation as such. This may explain why researchers looking for simple, independent, resource-based, indices of 'environmental' influence on IQ have been disappointed (Bradley, 1994; Turkheimer & Waldron, 2000). Numerous 'environmental' correlates of IQ have been frequently reported: demographic factors like social class, ethnic group, father's occupation and level of education, urban versus rural dwelling, and family characteristics like mother's style of parenting (responsiveness, amount of verbal and other stimulation in the home), and so on. But there is still little understanding of causal mechanisms linking these with IQ and cognitive development (Grissmer, Williamson, Kirby, & Berends,1998; Sternberg, 1995a; Wachs, 1992; Wozniak, 1993). This may also explain the (at best) mixed results of attempts to boost IQ and educational achievement (Mackintosh 1998).

In sum, ecological and sociohistorical theories view cognitive systems as having evolved for dealing with environmental change and unpredictability via developed knowledge representations, rather than as an all-purpose computational device. This means that cognitive processing and its variation is heavily knowledge-based, and, in humans, structured by engagement in specific cultural tools. IQ tests simply screen for individuals' uptake of one particular set of cultural/psychological tools, rather than their ability for complex cognition as such.

Special Affective Preparedness

Another tacit assumption of the computational model underlying IQ is that testees operate in a social and affective vacuum. However, humans also have complex values, beliefs, attitudes, motives, self-concepts and feelings, which make them more or less well prepared for specific testing situations and engagement with them. As might be expected, cognitive performance varies with degree of cognitive engagement with the task, and correlations between 'maximal' and 'typical' performance tests are generally weak (e.g. Goff & Ackerman, 1997). Martinez (2000) reviews research indicating that it is 'mindfulness', or degree of critical engagement with tasks, that determines success in them, failures being due to lack of persistence and premature self-termination, rather than lack of cognitive ability as such.

This, of course, is a broader use of the term 'affective' than is usual, implying a range of non-cognitive factors and dispositions that affect performance because of reduced engagement. What factors may influence subjects' engagement in cognitive tests? Some of these may be subsumed under the term 'academic orientation'. This includes such variables as parental interest and encouragement with school-related learning; their educational and occupational aspirations for their children; maternal involvement in play; provision of opportunities for learning; and so on. Correlations of .6–.7 between such variables and IQ have been regularly reported (Mackintosh, 1998; Suzuki & Valencia, 1997; Widlack & Perrucci, 1988).

Again, the causal import of such correlations cannot be interpreted in terms of simple 'independent effects'; they probably reflect subtle processes operating inside and outside families, with origins in parents' and children's interactions with the wider social structure. For example, cultural values operating in school and society can issue as negative social evaluations and systematic prejudice on some parents and children, which may depress their cognitive engagement. Thus, strong associations have been reported between early teacher expectations of their pupils and the latters' later educational attainments, even when initial attainments have been controlled for (Blatchford, Burke, Farquhar, Plewis, & Tizard, 1989). Rist (1970) said that this was due to pupils being typecast by superficial attributes like verbal skills, self-presentation and social class. Even superficial factors like language dialect (Giles, 1970), physical height (Gillis, 1982) and facial appearance (Langlois, 1986) affect teachers' and others' judgements of an individual's intelligence. Langlois (1986) indicates how such attributions fashion a self-concept in a self-fulfilling way:

> Children will learn to emit those behaviors that are consistent with the expectations, attitudes and behaviors of their parents and other socializing agents. When some children are treated as if they will be popular, friendly and smart and when other children are treated as if they will be unpopular, aggressive and not smart, these two groups of children may then come to fulfil their prophecies and come to behave 'appropriately'. (p. 40)

Schaffer (1996) argues that class/cultural experiences create specific 'belief systems' in parents that determine parental rearing practices and the course of child development (see also Wozniak, 1993). Sameroff, Sunter, Baracas, Zax and Greenspan (1987) reported an association between parental belief systems and IQ. Such belief systems include powerful self-evaluations of personal cognitive competence, or cognitive self-efficacy beliefs. Bandura (1997) defines self-efficacy as people's judgements of their capabilities for organizing and executing required courses of action. Such beliefs have been shown to be strongly associated with educational attainments and career aspirations, irrespective of actual abilities (e.g. Wood & Locke, 1987; for review see Schunk, 1995). As Bandura, Barbaranelli, Caprara and Pastorelli (1996) note:

> The findings of diverse lines of research reveal that efficacy beliefs exert considerable impact on human development and adaptation. . . . Such beliefs influence aspirations and strength of goal commitments, level of motivation and perseverance in the face of difficulties and setbacks, resilience to adversity, quality of analytical thinking, causal attributions for successes and failures, and vulnerability to stress and depression. (p. 1206)

Bandura et al. (1996) have also shown how cognitive self-efficacy beliefs acquired by parents in these ways are (socially) inherited by their children, resulting in significant depressions of expectations of personal performance in many intellectual tasks. Schunk (1995) reviews various studies showing how students acquire self-efficacy information vicariously from parents, teachers and peers, and indicate the importance of 'models' (presence or absence of others successfully tackling problems).

Obviously, children who receive doubts about their learning abilities from parents are more likely to avoid cognitive engagement in unusal problem solving. They may develop a syndrome known as 'learned helplessness'. Dweck and colleagues (e.g. Smiley & Dweck, 1994) have shown how conceptions of personal levels of intelligence either empower the individual or render him or her vulnerable to situational pressures. 'Helpless' children experience negative feelings, underestimate their past performance, predict poor future performance, and thus actually perform less well. 'Mastery-oriented' children have more positive feelings, galvanize their cognitive self-efficacy, focus on ways of improving performance and actually perform better in future. Even 'children as young as preschoolers display every aspect of the helpless pattern: self-blame, lowered expectations, negative emotions, lack of persistence and impaired strategies' (Smiley & Dweck, 1994, p. 262).

Determinants of such factors are likely to be strongly social-class-related (for review see Maddux, 1995). As Jenkins (1991) explains, on the basis of survey evidence, 'people of lower social class are likely to encounter more adverse experiences and fewer supportive experiences, and to be less in control of their environments' (p. 107). Moreover, working-class parents' low self-efficacy beliefs may explain why they tend to be more 'directive' in language interactions with children, which are less preparatory for IQ test performance (Martinez, 2000). This kind of 'affective' preparedness is, therefore, likely to constitute much of the variance in IQ scores and scholastic attainments: that is, another aspect of what IQ tests test.

Performance Preparedness

Reduced 'affective' preparedness will also reduce cognitive engagement with tests by reducing self-confidence and increasing anxiety in testing situations. As Maddux (1995) explains, 'Self-efficacy beliefs are powerful determinants of affective or emotional responses . . . that can then influence cognition and action' (p. 14). Thus, Stankov and Crawford (1997) found an association between self-confidence ratings and performance on the Raven's Progressive Matrices. Results of Stankov (2000) indicate a 'strong self-confidence factor' described as 'on the borderline between personality and intelligence' (p. 138) in IQ test performance, seemingly affecting attentional mechanisms and error-proneness.

Kirsch (1995) suggests that self-efficacy beliefs manifest as anxiety in threatening conditions, and therefore lead to avoidance behaviour. Premature self-termination on cognitive tests may be one aspect of this (Baron, 1985). There is a well-known association between high test anxiety (HTA) and IQ test performance. In his meta-analysis of 562 studies, Hembree (1988) found that HTA subjects experience more 'encoding difficulty', and more cognitive interference in the testing situation. Mikulincer (1994) has reviewed a number of studies showing strong relationships between anxiety and performance on various cognitive tasks stemming from 'impairment of several aspects of information-processing, including problem solving, attention, memory and categorization' (p. 166). Zeidner (1995) likewise reviews a large number of studies reporting associations between test anxiety and cognitive test performance. Raven, Raven and Court (1993, p. G14) note how ‘fatigue, ill health and stress’ affect speed and accuracy on the RPM.

As mentioned by Bandura et al. (1996), these effects can be (non-genetically) transgenerational. Parents with low expectations of themselves have low expectations of their children, so affecting children’s reactivity and anxiety in less familiar circumstances, in a way continued across generations. Some of this may entail a (non-genetic) biological as well as a social route. For example, in rodents it has been shown how stress reactivity, and the regulation of genetic expression in the brain associated with it, can be transmitted from one generation to the next through maternal behaviour with offspring (Francis, Diorio, Liu, & Meaney, 1999). This reactivity is likely to disrupt attention and may explain the reported correlations between habituation to novel stimuli in infancy and later IQ (McCall & Carriger, 1993), as well as more general IQ variance.

The idea that at least some of the variance in IQ scores is created not only by current individual differences, but also, in effect, by individuals’ socio-history across generations may shed further light on the nature of the IQ test as a descriptive instrument. In any case, it appears quite likely that affect-related performance preparedness may constitute another major source of variance in IQ variance—another aspect of what IQ tests test.

IQ as a Surrogate for Social Class Affiliation

These studies all point to a strong chain of causation from sociocognitive-affective preparedness to differences in performance on cognitive tasks and educational attainments. This nexus of factors is strongly social class/ethnic-group related (see also Fischer et al., 1993). It suggests more tractable sources of IQ variance than the anonymous *g*, or general mental ability, favoured by IQ theorists.

This idea is supported by what has become known as the ‘Flynn effect’, after surveys showing the steady rise in average IQ scores in many populations over recent decades (Flynn, 1998; Neisser, 1998). This is a continuing puzzle to *g* theorists because it could not be due to genetic changes over the short period in question. And as Flynn (1998) notes, ‘IQ gains have not been accompanied by an escalation of real world cognitive skills . . . an evolution from widespread retardation to normalcy, or from normalcy to widespread giftedness’ (p. 61). These leaps in average IQs do, however, correspond to the demographic swelling of the middle classes over the period in question. This means greater exposure to middle-class cultural tools, improved self-esteem and self-confidence, as well as more specific symbolic and technological procedures (Greenfield, 1998). This explanation for the Flynn effect is also supported by the fact that it is more prominent with non-verbal tests like the RPM than with verbal tests (Flynn, 1998).

The idea that IQ is, in effect, simply a measure of social class (here viewed as degree of sociocognitive-affective preparedness for IQ tests) is sometimes dismissed with reference to other correlations between IQ and socioeconomic status (SES) (Brody, 1997; Mackintosh, 1998). For example, Jensen (1998) says that IQ is only moderately associated with actual SES ratings, with typical correlations of only .3–.4 for children, and that IQ is a better predictor of adult social status than parents' SES (therefore, that IQ is measuring something more than mere SES). As Jensen (1998) puts it, 'SES is an effect of IQ rather than a cause' (p. 491).

The problem with this line of argument is that of assuming that parents' SES, defined almost always as current occupation, income or level of education, is itself a precision measure of social class. On the contrary, as Mills (1995) explains, 'The economic and social factors are one thing: psychological feelings may or may not be associated with them in expected ways' (p. 208). As Suzuki and Valencia (1997) concur, what parents 'do' is more important than 'what they are'. Social class is a compound of the cultural tools (knowledge and cognitive and psycholingustic structures) individuals are exposed to; and beliefs, values, academic orientations, self-efficacy beliefs, and so on. Such factors are consistently better predictors of IQ and attainments than is SES (Martinez, 2000). Mackintosh (1998) reviews a number of studies showing that parental 'attitudes' to education and achievement are better predictors of children's IQ than is SES.

Yet parents in the same occupation (SES) may vary enormously on such criteria. For example, social mobility is great, so that for many parents, current SES may be a relatively recent status. In the (British) National Child Development Study (Fogelman, 1983) 16 per cent of fathers had changed their SES in the short period from their children being 7 to 11 years of age. Total SES mobility rates from one generation to the next of between 40 and 70 per cent have been reported in a variety of other studies in Europe and the USA (Erikson & Goldthorpe, 1994). Analyses of class ideologies show strong limits to class emulation among mobile groups, and that children from the same SES may have widely different patterns of aspirations, attitudes, responses to social challenges and acceptance of the dominant achievement ideology (Bersman & Vidich, 1995; MacLeod, 1994). Moreover, marriages are sometimes cross-SES, whilst SES mobility appears to be much more restricted among women compared with men. The significance of this is that whilst it is usually the father's occupation that is taken as the family SES, it is mothers' belief systems that are most related to the direction of their children's cognitive development.

Finally, the idea that SES accurately reflects a uniform 'family environment' relevant for IQ development is discredited by the fact that children in the same family can be treated very differently in ways affecting sociocognitive-affective preparedness. For example, it is well known that birth order is one of the most important predictors of school achievement

(Mackintosh, 1998), academic orientation seemingly impinging more on early born, compared with later, children. Sigel et al. (1993) also found that children from close-spaced, compared with widely spaced, families performed quite differently on cognitive tasks.

All these social dynamics will render SES only a weak index of the social variables most related to IQ variance, and militate against a high IQ/SES correlation for children. It seems reasonable to suggest instead that IQ, as a ‘middle-class proximity quotient’, is itself a much better psychological index of social class than is SES.

Having now made the case for a nexus of sociocognitive-affective variables as the true source of the variance tapped by IQ—in contrast to an anonymous general cognitive factor, or *g*—I will now attempt to show how that nexus may explain the various other correlates of IQ usually evoked in support of *g*.

Correlation with Educational Attainment

Many people are convinced that IQ tests are valid tests of ‘intelligence’ because test scores predict school achievement moderately well (correlations of around .5; Brody, 1997). This galvanizes the belief that IQ tests are measuring an independent source of cognitive variance causal to variation in school attainment. But there are problems with this view. There have been criticisms about making causal inferences from what is only correlational data (e.g. Howe, 1998). In any case, at least part of this covariation is ‘built-in’ by the method of test construction. As Thorndike and Hagen (1969) explained, ‘From the very way in which the tests were assembled it could hardly be otherwise’ (p. 325). Indeed, Binet and Simon admitted that at least four sources of variance are conflated in their test: ‘the intelligence pure and simple’; ‘extra-scholastic acquisitions’; ‘scholastic acquisitions’; and ‘acquisitions relative to language’, which depend ‘partly on the school and partly on the family circumstances’ (quoted by Zenderland, 1998, p. 130). That IQ is at least partly a product (rather than a cause) of school-related learning is also shown by the fact that length of schooling is itself a substantial predictor of later IQ (Ceci & Williams, 1997; Wahlsten, 1997).

The important question, of course, is whether the association can even partially be due to an independent agent, ‘the intelligence pure and simple’. Again, this remains a hypothesis supported by little direct evidence (Howe, 1998). Mackintosh (1998) maintains that the association must involve ‘something’ else, because ‘culture-free’ tests like the RPM also predict school attainments. However, as argued above, such tests may be more closely related to school-relevant cognitive preparedness factors than to ordinary verbal items. In other words, the need to posit an independent, hypothetical and undescribed source of variance (the anonymous *g*) to explain the IQ/school-attainment correlation remains questionable.

Why Scores on Different Tests Intercorrelate

The way that performances of individuals on different tests tend to intercorrelate has, since Spearman (1927), been taken as the strongest evidence of a unitary cognitive variance factor, or *g*, underlying IQ scores. Jensen (1998, p. 106) says that a common factor always emerges 'among all mental tests however diverse', and argues that this proves the *g* thesis. The point is reinforced (but hotly debated) by numerous factor-analytic studies, and also by the fact that the 'general factor' scores distilled from individual performances correlate more highly with other assumed indices of intelligence, such as educational attainments, than do the raw scores themselves (Jensen, 1998; for brief review see Carroll, 1997).

Many problems surround this view. First, the data are, again, statistical, with causal agency remaining hypothetical and undescribed. As Herrnstein and Murray (1994) warned, 'The evidence for a general factor was . . . circumstantial, based on statistical analysis rather than direct observation. Its reality therefore was, and remains, arguable' (p. 72). Glymour (1997) has presented thorough arguments about the many unsupported statistical assumptions that enter into factor analysis. A correlation between test scores does not necessarily mean that they are measuring the same thing. As Raven et al. (1993) put it, 'height and weight are correlated to much the same extent as "academic abilities"—yet height and weight are clearly not the same thing' (p. G8). Nor does it mean that any common source of variance between mental tests is necessarily cognitive. For example, Spearman's original hypothesis about *g* was based on intercorrelations between school attainments on various subjects. But, as noted above, surveys consistently show that the biggest source of variance in school attainments is parental 'academic orientation'. A child who is being motivated or 'pushed' by parents will tend to put in above-average effort in all subjects, and conversely for those who are poorly motivated. In addition it is well known that other explanations of Spearman's (and more recent factor-analytic) data are possible without recourse to a 'general factor' (e.g. Brown & Thomson, 1940; see also Sternberg, 1999).

Perhaps the more significant problem is that IQ test scores—in contrast to objective scientific measurements—are not direct, unadulterated renditions of naturally occurring variation. In principle, all standardized psychometric tests have been subjected to considerable construction engineering on the basis of common preconceptions about population variance in intelligence. So what is common may be 'built in' by the methods of test construction, rather than captured, as it were, from nature. Jensen (1980) himself has noted how in 'the best standardized tests . . . every item is carefully edited and selected on the basis of technical procedures known as "item analysis", based on tryouts of the items on . . . the test's target population' (p. 145). Terman (1916) described how items on the Stanford–Binet test were painstakingly selected or deselected according to their agreement with performance on the scale as a whole, so creating the appearance of measuring 'one massive common factor' (Butcher, 1968, p. 221). Most other tests have followed the Stanford–Binet in this regard (and, indeed are usually 'validated' by their level of agreement with it; Anastasi, 1990).

Further construction engineering compounds this artificial convergence to a general factor. For example, items are almost universally devised in the first place by people from a very narrow cultural background on the basis of 'face validity', with more or less intuitive reference to common criteria, such as school-type knowledge. This also applies to non-verbal items, such as the Raven's matrices, as mentioned above. Cole (1999), indeed, describes the leap of psychological decentring that would be required on the part of test constructors to overcome this cultural bias. Then items are further selected/deselected to ensure test properties further consistent with a simple quantitative (biometric) trait, including: a normal distribution of population scores; a linear age-wise increase in scores; no sex differences; and so on. Out of this vast technology, it is, perhaps, not surprising that a common factor emerges, though it may be one in test designers' presuppositions rather than one in nature.

It seems at least as plausible to argue that the 'common factor' being measured in IQ tests is actually the nexus of sociocognitive and affective preparedness factors described above. So when *g* is statistically abstracted from test performances, it is simply purifying this nexus (partialling out extraneous sources of variance). This is why correlates of *g* with other subjectively assumed criteria of 'intelligence' tend to be higher than those of raw IQ scores.

Correlations with Occupational Status and Performance

Because IQ is associated with school achievement, which is associated with level of entry to the job market, there is an inevitable correlation ($r = \sim .5$) between IQ and occupational status (Neisser et al., 1996). Again, the extent to which this correlation can be causally attributed to an anonymous general cognitive factor, or *g*, is unknown.

The question that has created most debate, however, is whether there is an association between IQ scores and job performance. This has been difficult to establish because of uncertain suitability of tests, poor reliability of performance ratings, contradictory findings and, of course, selectivity of reporting and citation. (In addition, there is much use in reports of correlations 'corrected' for unreliability, which boosts them by as much as 50–100 per cent. These need to be treated with caution because a 'correction' only records the highest theoretical correlation, *given* such unreliable instruments, not necessarily the 'true' correlation. In any case, such correlations cannot, of course, automatically be interpreted as causal.)

The most-used indices of job performance are ratings of supervisors, with tests of work-related knowledge and judgements of work samples also used to some extent. Jensen (1980) has noted how supervisors can be very subjective, and use widely different criteria in making their judgements. In addition, 'age effects' and 'halo' effects have frequently been reported (e.g. Murphy & Balzer, 1986). In a frequently cited study by Schmidt, Hunter and Outerbridge (1986), supervisor ratings had only a correlation of .3 with subjects' job knowledge, and virtually zero correlations with actual samples of work performance. On these grounds alone reported correlations need to be treated with caution.

In fact correlations between IQ and estimates of job performance tend to be very low, averaging around .3 (Neisser et al., 1996). Wagner (1997) summarizes the few studies that have been done by pointing out that cognitive ability has little if any direct effect on job performance (near zero), with an indirect association via job knowledge ($r = \sim .3$). Raven et al. (1993, p. G11) say that the RPM accounts for no more than 10 per cent of the variance in occupational performance. As Wagner says, disentangling causal effects from these associations requires additional constructs.

One such construct may be the nexus of sociocognitive-affective preparedness factors described above. Consistent with this interpretation are reports that, when workers have been in the job for some time (and their confidence and anxiety levels have presumably improved), performance is completely unassociated with IQ (Hulin, Henry, & Noon, 1990). This ties in with the empirical demonstration that differences in IQ seem to be associated with learning of complex skills in the early period, but not thereafter (Ackerman, 1988; cf. Mackintosh 1998). Moreover, IQ scores do not seem to distinguish the more from the less successful groups (as judged by peers) in a wide range of occupations (reviewed by Ericcson & Charness, 1994; Mackintosh, 1998).

In sum, it seems likely that what genuine association there is between IQ and job performance can be more transparently attributed to the sociocognitive-affective preparedness factors mentioned above than to the uncharacterized *g*.

Correlations with Reaction Times/Iinspection Times

A number of studies over the last twenty years or so have reported correlations between IQ and performance on 'elementary cognitive tasks' (ECTs). These include simple and choice reaction times (RTs; based on whether the subject has to respond to one or more than one response target) and 'inspection time' (IT; based on the time needed to discriminate accurately between two perceptual values, such as the lengths of two lines, presented for very short intervals, usually a few milliseconds). The argument is that these simple tasks engage basic neurophysiological processes independently of background knowledge or other acquired cognitive structures. It is variance in such processes that Jensen (1998) sees as the very roots of *g* (IQ variance) itself.

Reported raw correlations vary greatly, but average around –.3 (faster responses associated with higher IQ; Deary & Stough, 1996; Neisser et al., 1996), though, here again, there appears to be a strong tendency for 'corrected' rather than raw correlations to be cited. Objections to these reports have been robust. Some have involved general criticism of the idea that variation in simple, peripheral processes can even theoretically be important to variation in 'intelligence'. Sub-human animals like rhesus monkeys are known to have faster RTs than have humans (Washburn & Rumbaugh, 1997). And if such a source of variance is critical to intelligence, then we would expect Grand Prix racing drivers, air force fighter pilots, and the like, to be the most intelligent humans (or highest in *g*). In addition, it has been suggested that an 'intelligence' in which individuals can display superiority/inferiority before they are even conscious of the task to which it is being applied is a strange intelligence indeed (Bub, 2000).

On the other hand, many studies have now demonstrated the role of 'top-down', cognitive processes—influencing what Bub (2000) calls 'state of preparedness'—in all ECT responses. An optimum state of readiness for performing ECTs involves many factors like selective attention, the monitoring of expectancies, response preparation, filtering of extraneous thoughts and sensory distractions, modulation of internal states, and so forth (Bub, 2000). In concurrence with this view, Nettlebeck and Vita (1992) found large practice effects on an IT task, with corresponding diminution of IQ–IT correlation to 'negligible proportions'. Bors, Stokes, Forrin and Hodder (1999) report similar findings, and say that 'attentiveness is at least in part responsible for the IQ–IT correlation' (p. 111). Using a slightly different task, Burns, Nettlebeck and Cooper (2000) report little association between IT and 'fluid ability'. In addition, even individual ITs are not stable from one situation to another (Deary & Stough, 1996).

Again it would seem at least as plausible to argue that IQ/RT and IQ/IT correlations can be explained by the common role of sociocognitive-affective preparedness factors mentioned above. As we have seen, factors like test anxiety or self-confidence substantially affect IQ test performances. Other aspects of ECT data support this interpretation. For example, it is consistently found that the largest correlations with IQ are for intra-individual variation in RT (usually measured by individuals' standard deviations, or SDs, over numerous trials). As Jensen (1998) explains, 'It is a rare study indeed in which RTSD does not have a larger (negative) correlation with IQ than does RT itself. In other words, higher-IQ persons have more consistent RT's from trial to trial when performing an ECT' (p. 225).

In other words, low-IQ subjects regularly produce RTs equal to those of high-IQ subjects, but with less consistency over trials. This lack of consistency may well reflect poor self-confidence and high test anxiety and their effects on information processing, incursions of extraneous cognitions, sensory distractions and so on. This interpretation is reinforced by Jensen's 1998 (p. 224) report that RT significantly correlates (–.45) with Extraversion scores on the Eysenck Personality Inventory. Ratings on the EPI are related to high self-efficacy beliefs, self-confidence, freedom-from-anxiety and other aspects of 'emotional well-being' (Peterson, Maier, & Seligman, 1993). Again it seems reasonable to suggest that any common source of variance in IQ and ECTs originates in the sociocognitive-affective nexus described above.

IQ Does Not Measure the Ability for Complex Reasoning

Perhaps the strongest evidence against the notion that IQ tests measure general cognitive ability is that IQ scores appear to be unrelated to the ability to carry out complex problem solving in social and practical situations. As mentioned above, even non-verbal or 'performance' items such as Raven's matrices do not appear to be particularly complex, cognitively, compared with everyday social and practical problem situations. In a study of betting at a racecourse, Ceci and Liker (1986) found the punters' predictions of odds to be a sophisticated cognitive process, entailing values on up to 11 variables, involving non-linearities and complex interactions between them. They found individuals' accuracy at such predictions to be unrelated to IQ, which led the authors to conclude that, 'whatever it is that an IQ test measures, it is not the ability to engage in cognitively complex forms of reasoning' (p. 264).

Similar results have consistently emerged from studies using simulations of 'real-life' problem situations: for example, simulated factory-production or public-service systems with large numbers of variables, organized according to complex equations, mostly opaque to the subject who has to regulate the system (for review see Beckmann & Guthke, 1995). In a study by Putz-Osterloh and colleagues (see Putz-Osterloh & Lemme, 1987), subjects had to manage a miniature tailor's establishment, manipulating numerous cost and production variables in order to maximize profit. The investigators found no relationship between performance on this system and measured IQ, and concluded that 'real' situations involve a complexity of problem structure, and thus of reasoning, not reflected in IQ tests (cf. Funke, 1991). Similar results were found by Dorner and Wearing (1995) utilizing a simulated town operating over numerous input and outcome variables. However, another factor, 'self-confidence', was strongly related to performance in the complex problem solving. Sternberg (1999) refers to several studies with managers, salespersons and university professors, showing that IQ-type test scores have little if any correlation with performances on the kinds of tasks they regularly encounter in their jobs.

In a complex problem-solving task requiring the induction of deeply implicit rules (i.e. many variables, multiply interacting, and thus difficult to unpack in explicit linguistic form) in artificial grammars, Reber, Walkenfeld and Hernstadt (1991) found only small individual differences in performance between subjects and no association with IQ. However, performance on an 'explicit' task (completing letter sequences) did correlate with IQ. This, too, indicates how IQ is a measure of attunement to explicit—and invariably culturally mediated—processes, rather than complexity of cognition as such. This, of course, is precisely why Binet, in founding the modern IQ testing movement was forced to eschew misguided attempts to measure 'intelligence' directly and focus, instead, on their expressions in explicit (culturally grounded) tasks.

As shown in the Ceci and Liker (1986) and other studies, success in 'complex' problem solving depends on the utilization of higher-order interactions between predictor variables. Such 'deep' information structures—or 'hyperstructures' (Richardson, 1998)—would seem to be what characterizes complex, changeable (especially social) environments. Coping with them requires highly specific, and therefore developed, knowledge structures, rather than a 'general information-processing' capacity (Sternberg, 1995b).

Pitfalls of Using School Attainment as a Criterion of General Intelligence

As mentioned above, intelligence tests were first established as predictors of educational attainments, and their success in this regard has been the mainstay of the IQ testing movement. Although the circularity in such claims was also mentioned above, the unstated, and usually unquestioned, assumption behind them is that educational achievement is itself a 'measure' of intelligence. As Hunt (1997) put it, 'high school and college may be the most intellectually demanding stages of life' (p. 167).

However, analysis and critique of the schooling process draw this assumption into question. In their detailed observations, Edwards and Mercer (1987) found that school learning involves ‘ritual knowledge’ rather than real understanding. Cole (1990) suggests that school learning involves, in the main, large amounts of fragmented information to be committed to memory; basic communication and computation; and certain forms of knowledge classification. This limited form of learning is dictated by what Bruner (1985) calls ‘artificial “madeup” subjects’. Though useful for sorting children out on the basis of their motivation and persistence (themselves rooted in family backgrounds), they encourage only a very narrow kind of cognitive ability. They have little to do with knowledge as used in the practical world, or with knowledge as it is used by critical scholars in universities (Perret-Clermont & Bell, 1988). Accordingly, much research has shown that even high school and university students have difficulty applying their knowledge to practical problems only slightly different from the form in which they have been encountered in the curriculum (Gardner, 1993). Moreover, a clear relationship between individual proficiency in school learning and later proficiency in university and/or the world of work has always been difficult to establish (Cook, 1988; Peers & Johnston, 1994; Wagner, 1994).

Most school experience, in other words, seems to be set up to test children’s perseverance and learning confidence—in effect another test of their sociocognitive-affective preparedness rather than a ‘test’ of any general intelligence. As Raven (2000), puts it, ‘the so-called “educational system” fails to identify and develop most of the talents of most children. . . . [it] instead functions as a sociological system to legitimize and sustain huge differentials in wealth, well-being and power’ (p. 404). This conclusion is reinforced by studies indicating the cognitive complexity of many everyday social and practical activities (briefly described above), and the ease with which school subjects such as mathematics can be learned when embedded in socially relevant goals and purposes (Lave, 1994). In other words, using school achievement as a major criterion of intelligence is itself problematic: it appears to be based on the same specific kinds of sociocognitive-preparedness factors as the IQ test itself, rather than on intellectual ability as such.

Conclusion

The assertion that IQ measures human intelligence in any general sense, or that the source of variance in IQ scores is primarily cognitive in nature, remains unsubstantiated after decades of investigation. The hypothetical ‘general ability’, or *g*, remains as inscrutable as the ‘vital forces’ once thought to distinguish living from non-living things. I have argued, here, that

IQ variance arises from a nexus of sociognitive-affective factors that render individuals more or less well prepared for IQ tests in terms of specific facts and information, specific linguistic and cognitive structures, and the self-efficacy, self-confidence and levels of test anxiety that influence actual performance on them. These factors suggest more tractable causes underlying the range of correlations usually cited in support of *g*. They also help explain the 'Flynn effect', why 'culture-free' tests like the RPM are even more narrowly enculturated than most verbal items, and how IQ variation is created out of test technology with items that are, themselves, relatively undemanding in terms of cognitive complexity.

If the nature of IQ variance described here has any validity, then wide implications for psychology and beyond follow. At root the status of the IQ test as a scientific instrument, and IQ scores as scientific measures, is questioned. This reinforces the view of Michell (1999), who, after criticizing the technical assumptions that convert multiples of responses to a scalar quantity, concludes that 'attempts to measure psychological attributes do not yet stand as scientific results. They remain hypotheses, the truth of which has not been adequately tested' (p. 216).

Most debate about IQ seems to presuppose that IQ variance is basically that of a simple quantitative (biometric) trait (e.g. Neisser et al., 1996), and this may be the source of much confusion and controversy. The analysis presented here suggests it may be something other than that. This implies serious caution about many avenues of the interpretation and application of IQ scores. For example, inferring future potential for cognitive competence (e.g. in education or occupation) from a current IQ score may not only be misleading. Popularizations of IQ (*g*) theory create widely held belief systems and expectations about the abilities of self and others that may actually construct the grounds for that theory's self-fulfilment. This supports the view of Raven (2000) that it is in upholding a 'sociological system' that IQ tests appear to have become instrumental, rather than in the assessment and cultivation of human intelligence as such.

Being clearer about the true composition of IQ variance will have other implications. For example, it would indicate a quite different notion of the 'environment' of cognitive development. And those attempting to 'boost' IQ, as in compensatory education schemes, are urged to review what, exactly, would be boosted. Furthermore, if variance in human intelligence cannot be reduced to that of a simple biometric model, then this compromises a fundamental assumption of controversial efforts to estimate the 'heritability' of IQ. Accordingly, a highly intensive, expensive and highly publicized search for 'genes for *g*', now under way (Plomin, 1999), may be based on a fundamental misconception.

This analysis of what IQ tests test may encourage attempts to describe human intelligence 'beyond' IQ, such as those of Sternberg (1999) or Gardner (1983). Sternberg's theory of 'successful intelligence' (e.g. 1999) is intriguing, but still accepts the notion of *g*, largely on the basis of the correlational evidence criticized above. Gardner (e.g. 1983) rejects a '*g*' factor, but resorts to hypothetical, innate computational 'modules' to explain the rich variation in human intelligence. Both are, however, heavily based on variances in cognitive test scores, whereas a closer examination of the nature of the variance may suggest further theoretical possibilities.

Of course, the sociocognitive-affective preparedness hypothesis requires further investigation, and new lines of research are suggested by it. It is research, however, that will need to presuppose a more complex conception of human intelligence than a bald computational one; and a model of IQ variance more complex than a simple biometric one. Nearly a century ago Binet (see citations in Zenderland, 1998) was protesting about attempts to reduce intelligence, and its defects, to simple biophysical models, and urged, instead, the need for more complex *psychological* descriptions. It appears that this admonition needs to be heeded more urgently than ever.

References

Ackerman, P.L. (1988). Determinants of individual differences in skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General*, *117*, 299–318.

Anastasi, A. (1990). *Psychological testing.* New York: Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Bandura, A., Barbaranelli, C., Caprara, G.V., & Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child Development*, *67*, 1206–1222.

Baron, J. (1985). *Rationality and intelligence.* Cambridge: Cambridge University Press.

Beckmann, J.F., & Guthke, J. (1995). Complex problem solving, intelligence and learning ability. In P.A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective*. Hillsdale, NJ: Erlbaum.

Bersman, J., & Vidich, A.J. (1995). The new class-system and its lifestyles. In A.J. Vidich (Ed.), *The new middle classes*. London: Macmillan.

Blatchford, P., Burke, J., Farquhar, C., Plewis, I., & Tizard, B. (1989). Teacher expectations in infant school: Associations with attainment and progress. *British Journal of Educational Psychology*, *59*, 19–30.

Boring, E.G.(1923, June). Intelligence as the tests test it. *New Republic*, June, pp. 35–37.

Bors, D.A., Stokes, T.L., Forrin, B., & Hodder, S.L. (1999). Inspection time and intelligence: Practice, strategies and attention. *Intelligence*, *27*, 111–129.

Bradley, R.H. (1994). The HOME inventory: Review and reflections. In H.W. Reese (Ed.), *Advances in Child Development and Behavior, Vol. 25*. San Diego, CA: Academic Press.

Brody, N. (1997). Intelligence, schooling and society. *American Psychologist*, *52*, 1046–1050.

Brown, W., & Thomson, G. (1940). *The essentials of mental measurement* (4th ed.). Cambridge: Cambridge University Press.

Bruner, J.S. (1985). Vygotsky: A historical and contemporary perspective. In J.V. Wertsch (Ed.), *Culture, communication and cognition: Vygotskian perspectives.* Cambridge: Cambridge University Press.

Bub, D.N. (2000, April). Reflections on *g* and the brain: Book review of Jensen on Intelligence-*g*-factor [6 paragraphs]. *Psycoloquy* [On-line serial], 11(43). Available FTP: Hostname: princeton.edu Directory: pub/harnad/Psycoloquy/1999.volume 11 File: psycoloquy.00.11.043.intelligence-g-factor.42.bub

Burns, N.R., Nettlebeck, T., & Cooper, C.J. (2000). Inspection time correlates with general speed of processing but not with fluid ability. *Intelligence*, 27, 37–44.

Butcher, H.J. (1968). *Human intelligence.* London: Methuen.

Carey, S. (1988). Are children fundamentally different kinds of thinkers and learners than adults? In K. Richardson & S. Sheldon (Eds.), *Cognitive development to adolescence.* Hove: Erlbaum.

Carpenter, P.A., Just, M.A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97, 404–431.

Carroll, J.B. (1997). Theoretical and technical issues in identifying a factor of general intelligence. In B. Devlin, S.E. Feinberg, D.P. Resnick, & K. Roeder (Eds.), *Intelligence, genes and success.* New York: Copernicus.

Ceci, S.J., & Liker, J.K. (1986). A day at the races: A study of IQ, expertise, and cognitive complexity. *Journal of Experimental Psychology: General, 115,* 255–266.

Ceci, S.J., & Roazzi, A. (1994). Cognition in context. In R.J. Sternberg & R.K. Wagner (Eds.), *Mind in context: Interactionist perspectives on human intelligence.* Cambridge: Cambridge University Press.

Ceci, S.J., & Williams, M.W. (1997). Schooling, intelligence, and income. *American Psychologist, 52*, 1051–1058

Cheng, P.W., & Holyoak, K.J. (1985). Pragmatic reasoning schemes. *Cognitive Psychology, 17,* 391–416.

Chi, M., & Glaser, R. (1985). Problem solving ability. In R.J. Sternberg (Ed.). *Human abilities.* New York: Freeman.

Cole, M. (1990). Cognitive development and formal schooling: The evidence from cross-cultural research. In L.C. Moll (Ed.), *Vygotsky and education: Instructional implications and applications of sociohistorical psychology.* Cambridge: Cambridge University Press.

Cole, M. (1999). Culture-free versus culture-based measures of cognition. In R.J. Sternberg (Ed.), *The nature of cognition.* Cambridge, MA: MIT Press.

Cole, M., Engeström, Y., & Vasquez, O. (Eds.). (1997). *Mind, culture and activity: Seminal papers from the Laboratory of Comparative Human Cognition.* Cambridge: Cambridge University Press.

Cook, M. (1988). *Personnel selection and productivity.* Chichester: Wiley.

Danziger, K. (1990). *Constructing the subject.* Cambridge: Cambridge University Press.

Deary, I.J., & Stough, C. (1996). Intelligence and inspection time. *American Psychologist*, *51*, 599–608.

Detterman, D.K., & Sternberg, R.J. (Eds.). (1986), *What is intelligence? Contemporary viewpoints on its nature and definition.* Norwood, NJ: Ablex.

Devlin, B., Feinberg, S.E., Resnick, D.P., & Roeder, K. (Eds.). (1997). *Intelligence, genes and success*. New York: Copernicus.

Donald, M. (1991). *Origins of the modern mind*. Cambridge, MA: Harvard University Press.

Dorner, D., & Wearing, A.J. (1995). Complex problem solving: Toward a (computer simulated) theory. In P.A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective.* Hillsdale, NJ: Erlbaum.

Edwards, D., & Mercer, N.M. (1987). *Common knowledge.* London: Methuen.

Ericsson, K.A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, *49*, 725–747.

Erikson, R., & Goldthorpe, J.H. (1994). Trends in class mobility: The post-war European tradition. In D.B. Grunsky (Ed.), *Social stratification.* Oxford: Westview.

Fischer, K.W., Bullock, D.H., Rotenberg, E.J., & Raya, P. (1993). In R.H. Wozniak & K.W. Fischer (Eds.), *Development in context: Acting and thinking in specific environments.* Hove: Erlbaum.

Flynn, J.R. (1998). IQ gains over time: Towards finding the causes. In U. Neisser (Ed.), *The rising curve*. Washington, DC: American Psychological Association.

Fogelman, K. (Ed.). (1983). *Growing up in Great Britain.* London: Macmillan.

Francis, D., Diorio, J., Liu, D., & Meaney, M.J. (1999). Nongenetic transmission across generations of maternal behaviour and stress response in the rat. *Science*, *286*, 1155–1157.

Franklin, A.F. (1997). Sociolinguistic structure of word lists and ethnic group differences in categorical recall. In M. Cole, Y. Engeström, & O. Vasquez (Eds.), *Mind, culture and activity: Seminal papers from the Laboratory of Comparative Human Cognition.* Cambridge: Cambridge University Press.

Funke, J. (1991). Solving complex problems. In R.J. Sternberg & P.J. Frensch (Eds.), *Complex problem solving: Mechanisms and processes.* Hillsdale, NJ: Erlbaum.

Galton, F. (1883). *Inquiry into human faculty and its development.* London: Macmillan.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

Gardner, H. (1993). *The unschooled mind.* London: Fontana.

Giles, H. (1970). Evaluative reaction to accents. *Educational Review*, *23*, 211–227.

Gillis, J.S. (1982). *Too tall, too small*. Champaign, IL: Institute for Personality and Ability Testing.

Glymour, C. (1997). Social statistics and genuine inquiry: Reflections on *The bell curve*. In B. Devlin, S.E. Feinberg, D.P. Resnick, & K. Roeder (Eds.), *Intelligence, genes and success.* New York: Copernicus.

Goddard, H.H. (1911). Two thousand children tested by the Binet scale. *Journal of Education*, *74*, 99.

Goff, M., & Ackerman, P.L. (1997). Personality–intelligence relations: Assessment of typical intellectual engagement. *Journal of Educational Psychology*, *84*, 537–552.

Goswami, U. (1996). Analogical reasoning and cognitive development. *Advances in Child Development and Behavior*, *26*, 92–139.

Gottfredson, L.S. (1998). The general intelligence factor. *Scientific American*, *9*, 24–29.

Greenfield, P.M. (1998). The cultural evolution of IQ. In U.Neisser (Ed.) *The Rising Curve.* Washington, DC: American Psychological Association.

Grissmer, D.W., Williamson, S., Kirby, S.N., & Berends, M. (1998). Exploring the rapid rise in Black achievement scores in the United States (1970–1990). In U. Neisser (Ed.), *The rising curve*. Washington, DC: American Psychological Association.

Hart, B., & Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Brookes.

Heath, S.B. (1982). What no bedtime story means: Narrative skills at home and school. *Language in Society*, *11*, 49–76.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, 58, 47–77.

Herrnstein, R., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life.* New York: Free Press.

Howe, M. (1998) *IQ in question*. London: Sage.

Hulin, C.L., Henry, R.A., & Noon, S.L. (1990). Adding a dimension: Time as a factor in the generalizability of predictive relationships. *Psychological Bulletin*, *107*, 328–340.

Hunt, E. (1997). The concept and utility of intelligence. In B. Devlin, S.E. Feinberg, D.P. Resnick, & K. Roeder (Eds.), *Intelligence, genes and success.* New York: Copernicus.

Jenkins, R. (1991). Demographic aspects of stress. In C.L. Cooper & R. Payne (Eds.), *Personality and stress: Individual differences in the stress process.* London: Wiley.

Jensen, A.R. (1980). *Bias in mental testing.* New York: Free Press.

Jensen, A.R. (1998). *The g-factor.* Westport, CT: Praeger.

Keating D.P., & MacLean, D.J. (1987). Cognitive processing, cognitive ability, and development: A reconsideration. In P.A. Vernon (Ed.), *Speed of information processing and intelligence.* New York: Ablex.

Keil, F. (1988). On the structure-dependent nature of stages of cognitive development. In K. Richardson & S. Sheldon (Eds.), *Cognitive development to adolescence.* Hove: Erlbaum.

Kesner, R.P., & Olson, D.S. (1990). *Neurobiology of comparative cognition.* Hove: Erlbaum.

Kirsch, I. (1995). Self-efficacy and outcome expectancy. In J.E. Maddux (Ed.), *Self-efficacy, adaptation and adjustment: Theory research and application.* New York: Plenum.

Langlois, J.H. (1986). From the eye of the beholder to behavioral reality: Development of social behaviors and social relations as a function of physical attractiveness. In C.P. Herman, M. Zanna, & E.T. Higgins (Eds.), *Physical appearances, stigma and social behaviour.* Hillsdale, NJ: Erlbaum.

Lave, J. (1994). The culture of acquisition and the practice of understanding. In J.W. Stigler, R.H. Shweder, & G. Herdt (Eds.), *Cultural psychology.* Cambridge: Cambridge University Press.

Luria, A.R. (1976). *Cognitive development: Its cultural and social foundations.* Cambridge, MA: Harvard University Press.

Mackintosh, N.J. (1998). *IQ and human intelligence.* Oxford: Blackwell.

MacLeod, J. (1994). Ain't not makin' it: Levelled aspirations in a low-income neighbourhood. In D.B. Grunsky (Ed.), *Social stratification.* Oxford: Westview.

Maddux, J.E. (Ed.). (1995). *Self-efficacy, adaptation and adjustment: Theory research and application.* New York: Plenum.

Martinez, M.E. (2000). *Education as the cultivation of intelligence.* Mahwah, NJ: Erlbaum.

McCall, R.B., & Carriger, M.S. (1993). A meta-analysis of infant habituation and recognition memory performance as predictors of later IQ. *Child Development*, *64*, 57–79.

Michell, J. (1999). *Measurement in psychology.* Cambridge: Cambridge University Press.

Mikulincer, M. (1994). *Human learned helplessness: A coping perspective.* New York: Plenum.

Mills, C.W. (1995). The new middle class. In A.J. Vidich (Ed.), *The new middle classes.* London: Macmillan.

Miller, G.A. (1962). *Psychology: The science of mental life.* Harmondsworth: Penguin.

Murphy, K.R., & Balzer, W.K. (1986). Systematic distortions in memory-based behaviour ratings and performance evaluations. *Journal of Applied Psychology, 71*, 49–59.

Neisser, U. (Ed.). (1998). *The rising curve.* Washington, DC: American Psychological Association.

Neisser, U., Boodoo, G., Bouchard, T.J., Jr., Boykin, A.W., Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C., Perloff, R., Sternberg, R.J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101.

Nettlebeck, T., & Vita, P. (1992). IT in two childhood cohorts. *British Journal of Psychology*, *10*, 189–197.

Olson, D.R. (1986). Intelligence and literacy: The relationships between intelligence and the technologies of representation and communication. In R.J. Sternberg & R.K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world.* Cambridge: Cambridge University Press.

Peers, I., & Johnston, M. (1994). Influence of learning context on the relationship between A-level attainment and final degree performance: A meta-analytic review. *British Journal of Educational Psychology*, *64*, 1–18.

Perret-Clermont, A.-N., & Bell, N. (1988). Learning processes in social and instructional interactions. In K. Richardson & S. Sheldon (Eds.), *Cognitive development to adolescence.* Hove: Erlbaum.

Peterson, C., Maier, S.F., & Seligman, M.E.P. (1993). *Learned helplessness.* Oxford: Oxford University Press.

Plomin, R. (1999). Genetics and general cognitive ability. *Nature*, 402 (Supp.), C25–29.

Putz-Osterloh, W., & Lemme, M. (1987). Knowledge and its intelligent application in problem solving. *The German Journal of Psychology, 11*, 286–303.

Raven, J. (2000). Ethical dilemmas. *The Psychologist*, 13, 404–406.

Raven, J., Raven, J.C., & Court, J.H. (1993). *Manual for Raven's Progressive Matrices and Vocabulary Scales, Section 1.* Oxford: Oxford Psychologists Press.

Reber, A.S., Walkenfeld, F.F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *17*, 888–896.

Richardson, K. (1991). Reasoning with Raven—in and out of context. *British Journal of Educational Psychology*, *61*, 129–138.

Richardson, K. (1996). Putting Raven into context. *British Journal of Educational Psychology*, *66*, 535–538.

Richardson, K. (1998). *Models of cognitive development*. Hove: Psychology Press.
Richardson, K., & Webster, D. (1996). Analogical reasoning in context: A research note. *British Journal of Educational Psychology, 66*, 23–32.
Rist, R. (1970). Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education. *Harvard Education Review, 40*, 411–451.
Rogoff, B. (1993). Children's guided participation and participatory appropriation in sociocultural activity. In R.H. Wozniak & K.W. Fischer (Eds.), *Development in context: Acting and thinking in specific environments.* Hove: Erlbaum.
Sameroff, A.J., Sunter, R., Baracas, B., Zax, M., & Greenspan, S. (1987). IQ scores of 4-year-old children: Social environmental risk factors. *Pediatrics, 79*, 343–350.
Schaffer, H.R. (1996). *Social development*. Oxford: Blackwell.
Schmidt, F.L., Hunter, J.E., & Outerbridge, A.N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisor ratings. *Journal of Applied Psychology, 71*, 432–439.
Schunk, D.M. (1995). Self-efficacy and education. In J.E. Maddux (Ed.), *Self-efficacy, adaptation and adjustment*. New York: Plenum.
Sigel, I.E., Stinson, E.T., & Kim, M.-I. (1993). Socialization of cognition: The distancing model. In R.H. Wozniak & K.W. Fischer (Eds.), *Development in context: Acting and thinking in specific environments.* Hove: Erlbaum.
Smiley, P.A., & Dweck, C.S. (1994). Individual differences in achievement goals among young children. *Child Development, 65*, 1723–1743.
Spearman, C. (1927). *The abilities of man.* London: Macmillan.
Stankov, L. (2000). Complexity, metacognition and fluid intelligence. *Intelligence, 28*, 1221–143.
Stankov, L., & Crawford, J.D. (1997). Self-confidence and performance on tests of cognitive ability. *Intelligence, 25*, 93–109.
Sternberg, R.J. (1995a). For whom the bell curve tolls: A review of *The bell curve. Psychological Science, 5*, 257–261.
Sternberg, R.J. (1995b). Expertise in complex problem solving: A comparison of alternative perspectives. In P.A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective.* Hillsdale, NJ: Erlbaum.
Sternberg, R.J. (1999). Successful intelligence: Finding a balance. *Trends in Cognitive Sciences, 3*, 436–442.
Sternberg, R.J., & Berg, C.A. (1986). Quantitative integration: Definitions of intelligence: A comparison of the 1921 and 1986 symposia. In D.K. Detterman & R.J. Sternberg (Eds.), *What is intelligence? Contemporary viewpoints on its nature and definition.* Norwood, NJ: Ablex.
Suzuki, L.A., & Valencia, R.R. (1997). Race-ethnicity and measured intelligence. *American Psychologist, 52*, 1147–1150.
Terman, L. (1916). *The measurement of intelligence*. Boston, MA: Houghtom Mifflin.
Thorndike, R.L., & Hagen, E.P. (1969). *Measurement and evaluation in psychology and education*. New York: Wiley.
Turkheimer, E., & Waldron, M. (2000). Nonshared environment: A theoretical, methodological and quantitative review. *Psychological Bulletin, 126*, 78–108.
Vygotsky, L.S. (1988). The genesis of higher mental functions. In K. Richardson & S. Sheldon (Eds.), *Cognitive development to adolescence.* Hove: Erlbaum.
Wachs, T.D. (1992). *The nature of nurture*. London: Sage.

Wagner, R.K. (1994). Context counts: The case of cognitive ability testing for job selection. In R.J. Sternberg & R.K. Wagner (Eds.), *Mind in context: Interactive perspectives in human intelligence*. Cambridge: Cambridge University Press.
Wagner, R.K. (1997). Intelligence, training and employment. *American Psychologist*, *52*, 1059–1069.
Wahlsten, D. (1997). The malleability of intelligence is not constrained by heritability. In B. Devlin, S.E. Feinberg, D.P. Resnick, & K. Roeder (Eds.), *Intelligence, genes and success.* New York: Copernicus.
Washburn, D.A., & Rumbaugh, D.M. (1997). So why are we slower? *American Psychologist*, *52*, 1147–1150.
Wechsler, D. (1958). *The measurement and appraisal of adult intelligence.* Baltimore, MD: Williams & Wilkins.
Wertsch, J.V., del Río, P., & Álvarez, A. (Eds.). (1995). *Sociocultural studies of mind.* Cambridge: Cambridge University Press.
Widlack, P.A., & Perrucci, C.C. (1988). Family configuration, family interaction and intellectual attainment. *Journal of Marriage and the Family*, *50*, 33–44.
Wood, R.E., & Locke, E.A. (1987). The relation of self-efficacy and grade goals to academic performance. *Educational and Psychological Measurement*, *47*, 1013–1024.
Wozniak, R.H. (1993). Co-constructive metatheory for psychology: Implications for an analysis of families as specific contexts for development. In R.H. Wozniak & K.W. Fischer (Eds.), *Development in context: Acting and thinking in specific environments*. Hove: Erlbaum.
Wozniak, R.H., & Fischer, K.W. (Eds.). (1993). *Development in context: Acting and thinking in specific environments.* Hove: Erlbaum.
Zeidner, M. (1995). Personality trait correlates of intelligence. In D.H. Saklofske & M. Zeidner (Eds.), *International handbook of personality and intelligence.* New York: Plenum.
Zenderland, L. (1998). *Measuring minds: Henry Herbert Goddard and the origins of the American testing movement.* Cambridge: Cambridge University Press.

KEN RICHARDSON is formerly Senior Lecturer, and now Senior Visiting Research Fellow, in the Centre for Human Development & Learning at the Open University (UK). His main research and theoretical interests have been in cognitive development, with particular focus on evolutionary demands for a highly adaptable representational system, and how this reflects on issues surrounding human intelligence. Recent publications include *The Origins of Human Potential: Evolution, Development and Psychology* (Routledge, 1998), *Models of Cognitive Development* (Psychology Press, 1998) and *The Making of Intelligence* (Columbia University Press, 2000). ADDRESS: 13 Rectory Lane, Wolsingham, Co. Durham DL13 3AJ, UK. [email: k.richardson@open.ac.uk]