Intelligence: Knowns and Unknowns

Ulric Neisser (Chair) — *Emory University*
Gwyneth Boodoo — *Educational Testing Service, Princeton, New Jersey*
Thomas J. Bouchard, Jr. — *University of Minnesota, Minneapolis*
A. Wade Boykin — *Howard University*
Nathan Brody — *Wesleyan University*
Stephen J. Ceci — *Cornell University*
Diane F. Halpern — *California State University, San Bernardino*
John C. Loehlin — *University of Texas, Austin*
Robert Perloff — *University of Pittsburgh*
Robert J. Sternberg — *Yale University*
Susana Urbina — *University of North Florida*

In the fall of 1994, the publication of Herrnstein and Murray's book The Bell Curve *sparked a new round of debate about the meaning of intelligence test scores and the nature of intelligence. The debate was characterized by strong assertions as well as by strong feelings. Unfortunately, those assertions often revealed serious misunderstandings of what has (and has not) been demonstrated by scientific research in this field. Although a great deal is now known, the issues remain complex and in many cases still unresolved. Another unfortunate aspect of the debate was that many participants made little effort to distinguish scientific issues from political ones. Research findings were often assessed not so much on their merits or their scientific standing as on their supposed political implications. In such a climate, individuals who wish to make their own judgments find it hard to know what to believe.*

Reviewing the intelligence debate at its meeting of November 1994, the Board of Scientific Affairs (BSA) of the American Psychological Association (APA) concluded that there was urgent need for an authoritative report on these issues—one that all sides could use as a basis for discussion. Acting by unanimous vote, BSA established a Task Force charged with preparing such a report. Ulric Neisser, Professor of Psychology at Emory University and a member of BSA, was appointed Chair. The APA Board on the Advancement of Psychology in the Public Interest, which was consulted extensively during this process, nominated one member of the Task Force; the Committee on Psychological Tests and Assessment nominated another; a third was nominated by the Council of Representatives. Other members were chosen by an extended consultative process, with the aim of representing a broad range of expertise and opinion.

The Task Force met twice, in January and March of 1995. Between and after these meetings, drafts of the various sections were circulated, revised, and revised yet again. Disputes were resolved by discussion. As a result, the report presented here has the unanimous support of the entire Task Force.

1. Concepts of Intelligence

Individuals differ from one another in their ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought. Although these individual differences can be substantial, they are never entirely consistent: A given person's intellectual performance will vary on different occasions, in different domains, as judged by different criteria. Concepts of "intelligence" are attempts to clarify and organize this complex set of phenomena. Although considerable clarity has been achieved in some areas, no such conceptualization has yet answered all the important questions and none commands universal assent. Indeed, when two dozen prominent theorists were recently asked to define intelligence, they gave two dozen somewhat different definitions (Sternberg & Detterman, 1986). Such disagreements are not cause for dismay. Scientific research rarely begins with fully agreed definitions, though it may eventually lead to them.

This first section of our report reviews the approaches to intelligence that are currently influential, or that seem to be becoming so. Here (as in later sections) much of our discussion is devoted to the dominant *psychometric* approach, which has not only inspired the most research and attracted the most attention (up to this time) but is by far the most widely used in practical settings. Nevertheless, other points of view deserve serious consideration. Several current theorists argue that there are

This is a "Report of a Task Force Established by the American Psychological Association."

The Task Force appreciates the contributions of many members of the APA Board of Scientific Affairs (BSA) and the APA Board for the Advancement of Psychology in the Public Interest (BAPPI), who made helpful comments on a preliminary draft of this report. We also wish to acknowledge the indispensable logistical support of the APA Science Directorate during the preparation of the report itself.

Correspondence concerning the report should be addressed to Ulric Neisser, Department of Psychology, Emory University, Atlanta, GA 30322. Electronic mail may be sent via Internet to neisser@fs1.psy.emory.edu.

Vol. 51, No. 2, 77–101

many different "intelligences" (systems of abilities), only a few of which can be captured by standard psychometric tests. Others emphasize the role of culture, both in establishing different conceptions of intelligence and in influencing the acquisition of intellectual skills. Developmental psychologists, taking yet another direction, often focus more on the processes by which all children come to think intelligently than on measuring individual differences among them. There is also a new interest in the neural and biological bases of intelligence, a field of research that seems certain to expand in the next few years.

In this brief report, we cannot do full justice to even one such approach. Rather than trying to do so, we focus here on a limited and rather specific set of questions:

- What are the significant conceptualizations of intelligence at this time? (Section 1)
- What do intelligence test scores mean, what do they predict, and how well do they predict it? (Section 2)
- Why do individuals differ in intelligence, and especially in their scores on intelligence tests? Our discussion of these questions implicates both genetic factors (Section 3) and environmental factors (Section 4).
- Do various ethnic groups display different patterns of performance on intelligence tests, and if so what might explain those differences? (Section 5)
- What significant scientific issues are presently unresolved? (Section 6)

Public discussion of these issues has been especially vigorous since the 1994 publication of Herrnstein and Murray's *The Bell Curve,* a controversial volume which stimulated many equally controversial reviews and replies. Nevertheless, we do not directly enter that debate. Herrnstein and Murray (and many of their critics) have gone well beyond the scientific findings, making explicit recommendations on various aspects of public policy. Our concern here, however, is with science rather than policy. The charge to our Task Force was to prepare a dispassionate survey of the state of the art: to make clear what has been scientifically established, what is presently in dispute, and what is still unknown. In fulfilling that charge, the only recommendations we shall make are for further research and calmer debate.

The Psychometric Approach

Ever since Alfred Binet's great success in devising tests to distinguish mentally retarded children from those with behavior problems, psychometric instruments have played an important part in European and American life. Tests are used for many purposes, such as selection, diagnosis, and evaluation. Many of the most widely used tests are not intended to measure intelligence itself but some closely related construct: scholastic aptitude, school achievement, specific abilities, etc. Such tests are especially important for selection purposes. For preparatory school, it's the SSAT; for college, the SAT or ACT; for graduate school, the GRE; for medical school, the MCAT; for law school, the LSAT; for business school, the GMAT. Scores on intelligence-related tests matter, and the stakes can be high.

Intelligence tests. Tests of intelligence itself (in the psychometric sense) come in many forms. Some use only a single type of item or question; examples include the Peabody Picture Vocabulary Test (a measure of children's verbal intelligence) and Raven's Progressive Matrices (a nonverbal, untimed test that requires inductive reasoning about perceptual patterns). Although such instruments are useful for specific purposes, the more familiar measures of general intelligence—such as the Wechsler tests and the Stanford-Binet—include many different types of items, both verbal and nonverbal. Test-takers may be asked to give the meanings of words, to complete a series of pictures, to indicate which of several words does not belong with the others, and the like. Their performance can then be scored to yield several subscores as well as an overall score.

By convention, overall intelligence test scores are usually converted to a scale in which the mean is 100 and the standard deviation is 15. (The standard deviation is a measure of the variability of the distribution of scores.) Approximately 95% of the population has scores within two standard deviations of the mean, i.e., between 70 and 130. For historical reasons, the term "IQ" is often used to describe scores on tests of intelligence. It originally referred to an "Intelligence Quotient" that was formed by dividing a so-called mental age by a chronological age, but this procedure is no longer used.

Intercorrelations among tests. Individuals rarely perform equally well on all the different kinds of items included in a test of intelligence. One person may do relatively better on verbal than on spatial items, for example, while another may show the opposite pattern. Nevertheless, subtests measuring different abilities tend to be positively correlated: people who score high on one such subtest are likely to be above average on others as well. These complex patterns of correlation can be clarified by factor analysis, but the results of such analyses are often controversial themselves. Some theorists (e.g., Spearman, 1927) have emphasized the importance of a general factor, g, which represents what all the tests have in common; others (e.g., Thurstone, 1938) focus on more specific group factors such as memory, verbal comprehension, or number facility. As we shall see in Section 2, one common view today envisages something like a hierarchy of factors with g at the apex. But there is no full agreement on what g actually means: it has been described as a mere statistical regularity (Thomson, 1939), a kind of mental energy (Spearman, 1927), a generalized abstract reasoning ability (Gustafsson, 1984), or an index measure of neural processing speed (Reed & Jensen, 1992).

There have been many disputes over the utility of IQ and g. Some theorists are critical of the entire psychometric approach (e.g., Ceci, 1990; Gardner, 1983; Gould, 1978), while others regard it as firmly established (e.g., Carroll, 1993; Eysenck, 1973; Herrnstein & Murray, 1994; Jensen, 1972). The critics do not dispute the stability of test scores, nor the fact that they predict certain forms of achievement—especially school achievement—rather effectively (see Section 2). They do argue, however, that to base a concept of intelligence on test scores alone is to ignore many important aspects of mental ability. Some of those aspects are emphasized in other approaches reviewed below.

Multiple Forms of Intelligence

Gardner's theory. A relatively new approach is the theory of "multiple intelligences" proposed by Howard Gardner in his book *Frames of Mind* (1983). Gardner argues that our conceptions of intelligence should be informed not only by work with "normal" children and adults but also by studies of gifted persons (including so-called "savants"), of virtuosos and experts in various domains, of valued abilities in diverse cultures, and of individuals who have suffered selective forms of brain damage. These considerations have led him to include musical, bodily-kinesthetic, and various forms of personal intelligence in the scope of his theory along with more familiar linguistic, logical-mathematical, and spatial abilities. (Critics of the theory argue, however, that some of these are more appropriately described as special talents than as forms of "intelligence.")

In Gardner's view, the scope of psychometric tests includes only linguistic, logical, and some aspects of spatial intelligence; other forms have been almost entirely ignored. Even in the domains on which they are ostensibly focused, the paper-and-pencil format of most tests rules out many kinds of intelligent performance that matter a great deal in everyday life, such as giving an extemporaneous talk (linguistic) or being able to find one's way in a new town (spatial). While the stability and validity of performance tests in these new domains are not yet clear, Gardner's argument has attracted considerable interest among educators as well as psychologists.

Sternberg's theory. Robert Sternberg's (1985) triarchic theory proposes three fundamental aspects of intelligence—analytic, creative, and practical—of which only the first is measured to any significant extent by mainstream tests. His investigations suggest the need for a balance between analytic intelligence, on the one hand, and creative and especially practical intelligence on the other. The distinction between analytic (or "academic") and practical intelligence has also been made by others (e.g., Neisser, 1976). Analytic problems, of the type suitable for test construction, tend to (a) have been formulated by other people, (b) be clearly defined, (c) come with all the information needed to solve them, (d) have only a single right answer, which can be reached by only a single method, (e) be disembedded from ordinary experience, and (f) have little or no intrinsic interest. Practical problems, in contrast, tend to (a) require problem recognition and formulation, (b) be poorly defined, (c) require information seeking, (d) have various acceptable solutions, (e) be embedded in and require prior everyday experience, and (f) require motivation and personal involvement.

One important form of practical intelligence is *tacit knowledge,* defined by Sternberg and his collaborators as "action-oriented knowledge, acquired without direct help from others, that allows individuals to achieve goals they personally value" (Sternberg, Wagner, Williams, & Horvath, 1995, p. 916). Questionnaires designed to measure tacit knowledge have been developed for various domains, especially business management. In these questionnaires, the individual is presented with written descriptions of various work-related situations and asked to rank a number of options for dealing with each of them. Measured in this way, tacit knowledge is relatively independent of scores on intelligence tests; nevertheless it correlates significantly with various indices of job performance (Sternberg & Wagner, 1993; Sternberg et al., 1995). Although this work is not without its critics (Jensen, 1993; Schmidt & Hunter, 1993), the results to this point tend to support the distinction between analytic and practical intelligence.

Related findings. Other investigators have also demonstrated that practical intelligence can be relatively independent of school performance or scores on psychometric tests. Brazilian street children, for example, are quite capable of doing the math required for survival in their street business even though they have failed mathematics in school (Carraher, Carraher, & Schliemann, 1985). Similarly, women shoppers in California who had no difficulty in comparing product values at the supermarket were unable to carry out the same mathematical operations in paper-and-pencil tests (Lave, 1988). In a study of expertise in wagering on harness races, Ceci and Liker (1986) found that the reasoning of the most skilled handicappers was implicitly based on a complex interactive model with as many as seven variables. Nevertheless, individual handicappers' levels of performance were not correlated with their IQ scores. This means, as Ceci as put it, that "the assessment of the experts' intelligence on a standard IQ test was irrelevant in predicting the complexity of their thinking at the racetrack" (1990, p. 43).

Cultural Variation

It is very difficult to compare concepts of intelligence across cultures. English is not alone in having many words for different aspects of intellectual power and cognitive skill (*wise, sensible, smart, bright, clever, cunning . . .*); if another language has just as many, which of them shall we say corresponds to its speakers' "concept of intelligence"? The few attempts to examine this issue directly have typically found that, even within a given society, different cognitive characteristics are emphasized from one situation to another and from one subculture to another (Serpell, 1974; Super, 1983; Wober, 1974). These differences extend not just to conceptions of intelligence but also to what is considered adaptive or appropriate in a broader sense.

These issues have occasionally been addressed across subcultures and ethnic groups in America. In a study conducted in San Jose, California, Okagaki and Sternberg (1993) asked immigrant parents from Cambodia, Mexico, the Philippines, and Vietnam—as well as native-born Anglo-Americans and Mexican Americans—about their conceptions of child-rearing, appropriate teaching, and children's intelligence. Parents from all groups except Anglo-Americans indicated that such characteristics as motivation, social skills, and practical school skills were as or more important than cognitive characteristics for their conceptions of an intelligent first-grade child.

Heath (1983) found that different ethnic groups in North Carolina have different conceptions of intelligence. To be considered as intelligent or adaptive, one must excel in the skills valued by one's own group. One particularly interesting contrast was in the importance ascribed to verbal versus nonverbal communication skills—to saying things explicitly as opposed to using and understanding gestures and facial expressions. Note that while both these forms of communicative skill have their uses, they are not equally well represented in psychometric tests.

How testing is done can have different effects in different cultural groups. This can happen for many reasons. In one study, Serpell (1979) asked Zambian and English children to reproduce patterns in three different media: wire models, pencil and paper, or clay. The Zambian children excelled in the wire medium to which they were most accustomed, while the English children were best with pencil and paper. Both groups performed equally well with clay. As this example shows, differences in familiarity with test materials can produce marked differences in test results.

Developmental Progressions

Piaget's theory. The best-known developmentally-based conception of intelligence is certainly that of the Swiss psychologist Jean Piaget (1972). Unlike most of the theorists considered here, Piaget had relatively little interest in individual differences. Intelligence develops—in all children—through the continually shifting balance between the assimilation of new information into existing cognitive structures and the accommodation of those structures themselves to the new information. To index the development of intelligence in this sense, Piaget devised methods that are rather different from conventional tests. To assess the understanding of "conservation," for example (roughly, the principle that material quantity is not affected by mere changes of shape), children who have watched water being poured from a shallow to a tall beaker may be asked if there is now more water than before. (A positive answer would suggest that the child has not yet mastered the principle of conservation.) Piaget's tasks can be modified to serve as measures of individual differences; when this is done, they correlate fairly well with standard psychometric tests (for a review see Jensen, 1980).

Vygotsky's theory. The Russian psychologist Lev Vygotsky (1978) argued that all intellectual abilities are social in origin. Language and thought first appear in early interactions with parents, and continue to develop through contact with teachers and others. Traditional intelligence tests ignore what Vygotsky called the "zone of proximal development," i.e., the level of performance that a child might reach with appropriate help from a supportive adult. Such tests are "static," measuring only the intelligence that is already fully developed. "Dynamic" testing, in which the examiner provides guided and graded feedback, can go further to give some indication of the child's latent potential. These ideas are being developed and extended by a number of contemporary psychologists (Brown & French, 1979; Feuerstein, 1980; Pascual-Leone & Ijaz, 1989).

Biological Approaches

Some investigators have recently turned to the study of the brain as a basis for new ideas about what intelligence is and how to measure it. Many aspects of brain anatomy and physiology have been suggested as potentially relevant to intelligence: the arborization of cortical neurons (Ceci, 1990), cerebral glucose metabolism (Haier, 1993), evoked potentials (Caryl, 1994), nerve conduction velocity (Reed & Jensen, 1992), sex hormones (see Section 4), and still others (cf. Vernon, 1993). Advances in research methods, including new forms of brain imaging such as PET and MRI scans, will surely add to this list. In the not-too-distant future it may be possible to relate some aspects of test performance to specific characteristics of brain function.

This brief survey has revealed a wide range of contemporary conceptions of intelligence and of how it should be measured. The psychometric approach is the oldest and best established, but others also have much to contribute. We should be open to the possibility that our understanding of intelligence in the future will be rather different from what it is today.

2. Intelligence Tests and Their Correlates

The correlation coefficient, r, can be computed whenever the scores in a sample are paired in some way. Typically this is because each individual is measured twice: he or she takes the same test on two occasions, or takes two different tests, or has both a test score and some criterion measure such as grade point average or job performance. (In Section 3 we consider cases where the paired scores are those of two different individuals, such as twins or parent and child.) The value of r measures the degree of relationship between the two sets of scores in a convenient way, by assessing how well one of them (computationally it doesn't matter which one) could be used to predict the value of the other. Its sign indicates the direction of relationship: when r is negative, high scores on one measure predict low scores on the other. Its magnitude indicates the strength of the relationship. If $r = 0$, there is no relation at all; if r is 1 (or -1), one score can be used to predict the other score perfectly. Moreover, the square of r has a particular meaning in cases where we are concerned with predicting one variable from another. When $r = .50$, for example, r^2 is .25: this means (given certain linear assumptions) that 25% of the variance in one set of scores is predictable from the correlated values of the other set, while the remaining 75% is not.

Basic Characteristics of Test Scores

Stability. Intelligence test scores are fairly stable during development. When Jones and Bayley (1941) tested a sample of children annually throughout childhood and adolescence, for example, scores obtained at age 18 were correlated $r = .77$ with scores that had been obtained at age 6 and $r = .89$ with scores from age 12. When scores were averaged across several successive tests to remove short-term fluctuations, the correlations were even higher. The mean for ages 17 and 18 was correlated $r = .86$ with the mean for ages 5, 6, and 7, and $r = .96$ with the mean for ages 11, 12, and 13. (For comparable findings in a more recent study, see Moffitt, Caspi, Harkness, & Silva, 1993.) Nevertheless, IQ scores do change over time. In the same study (Jones & Bayley, 1941), the average change between age 12 and age 17 was 7.1 IQ points; some individuals changed as much as 18 points.

Is it possible to measure the intelligence of young infants in a similar way? Conventional tests of "infant intelligence" do not predict later test scores very well, but certain experimental measures of infant attention and memory—originally developed for other purposes—have turned out to be more successful. In the most common procedure, a particular visual pattern is shown to a baby over and over again. The experimenter records how long the infant subject looks at the pattern on each trial; these looks get shorter and shorter as the baby becomes "habituated" to it. The time required to reach a certain level of habituation, or the extent to which the baby now "prefers" (looks longer at) a new pattern, is regarded as a measure of some aspect of his or her information-processing capability.

These habituation-based measures, obtained from babies at ages ranging from three months to a year, are significantly correlated with the intelligence test scores of the same children when they get to be 2 or 4 or 6 years old (for reviews see Bornstein, 1989; Columbo, 1993; McCall & Garriger, 1993). A few studies have found such correlations even at ages 8 or 11 (Rose & Feldman, 1995). A recent meta-analysis, based on 31 different samples, estimates the average magnitude of the correlations at about $r = .36$ (McCall & Garriger, 1993). (The largest *r*s often appear in samples that include "at risk" infants.) It is possible that these habituation scores (and other similar measures of infant cognition) do indeed reflect real cognitive differences, perhaps in "speed of information processing" (Columbo, 1993). It is also possible, however, that—to a presently unknown extent—they reflect early differences in temperament or inhibition.

It is important to understand what remains stable and what changes in the development of intelligence. A child whose IQ score remains the same from age 6 to age 18 does not exhibit the same performance throughout that period. On the contrary, steady gains in general knowledge, vocabulary, reasoning ability, etc. will be apparent. What does *not* change is his or her score in comparison to that of other individuals of the same age. A six-year-old with an IQ of 100 is at the mean of six-year-olds; an 18-year-old with that score is at the mean of 18-year-olds.

Factors and g. As noted in Section 1, the patterns of intercorrelation among tests (i.e., among different kinds of items) are complex. Some pairs of tests are much more closely related than others, but all such correlations are typically positive and form what is called a "positive manifold." Spearman (1927) showed that in any such manifold, some portion of the variance of scores on each test can be mathematically attributed to a "general factor," or *g*. Given this analysis, the overall pattern of correlations can be roughly described as produced by individual differences in *g* plus differences in the specific abilities sampled by particular tests. In addition, however, there are usually patterns of intercorrelation among groups of tests. These commonalities, which played only a small role in Spearman's analysis, were emphasized by other theorists. Thurstone (1938), for example, proposed an analysis based primarily on the concept of group factors.

While some psychologists today still regard *g* as the most fundamental measure of intelligence (e.g., Jensen, 1980), others prefer to emphasize the distinctive profile of strengths and weaknesses present in each person's performance. A recently published review identifies over 70 different abilities that can be distinguished by currently available tests (Carroll, 1993). One way to represent this structure is in terms of a hierarchical arrangement with a general intelligence factor at the apex and various more specialized abilities arrayed below it. Such a summary merely acknowledges that performance levels on different tests are correlated; it is consistent with, but does not prove, the hypothesis that a common factor such as *g* underlies those correlations. Different specialized abilities might also be correlated for other reasons, such as the effects of education. Thus while the *g*-based factor hierarchy is the most widely accepted current view of the structure of abilities, some theorists regard it as misleading (Ceci, 1990). Moreover, as noted in Section 1, a wide range of human abilities—including many that seem to have intellectual components—are outside the domain of standard psychometric tests.

Tests as Predictors

School performance. Intelligence tests were originally devised by Alfred Binet to measure children's ability to succeed in school. They do in fact predict school performance fairly well: the correlation between IQ scores and grades is about .50. They also predict scores on school achievement tests, designed to measure knowledge of the curriculum. Note, however, that correlations of this magnitude account for only about 25% of the overall variance. Successful school learning depends on many personal characteristics other than intelligence, such as persistence, interest in school, and willingness to study. The encouragement for academic achievement that is received from peers, family, and teachers may also be important, together with more general cultural factors (see Section 5).

The relationship between test scores and school performance seems to be ubiquitous. Wherever it has been studied, children with high scores on tests of intelligence tend to learn more of what is taught in school than their lower-scoring peers. There may be styles of teaching and methods of instruction that will decrease or increase this correlation, but none that consistently eliminates it has yet been found (Cronbach & Snow, 1977).

What children learn in school depends not only on their individual abilities but also on teaching practices and on what is actually taught. Recent comparisons among pupils attending school in different countries have made this especially obvious. Children in Japan and China, for example, know a great deal more math than American children even though their intelligence test scores are quite similar (see Section 5). This difference may result from many factors, including cultural attitudes toward schooling as well as the sheer amount of time devoted to the study of mathematics and how that study is organized (Stevenson & Stigler, 1992). In principle it is quite possible to improve the school learning of American children—even very substantially—without changing their intelligence test scores at all.

Years of education. Some children stay in school longer than others; many go on to college and perhaps beyond. Two variables that can be measured as early as elementary school correlate with the total amount of education individuals will obtain: test scores and social class background. Correlations between IQ scores and total years of education are about .55, implying that differences in psychometric intelligence account for about 30% of the outcome variance. The correlations of years of education with social class background (as indexed by the occupation/education of a child's parents) are also positive, but somewhat lower.

There are a number of reasons why children with higher test scores tend to get more education. They are likely to get good grades, and to be encouraged by teachers and counselors; often they are placed in "college preparatory" classes, where they make friends who may also encourage them. In general, they are likely to find the process of education rewarding in a way that many low-scoring children do not (Rehberg & Rosenthal, 1978). These influences are not omnipotent: some high scoring children do drop out of school. Many personal and social characteristics other than psychometric intelligence determine academic success and interest, and social privilege may also play a role. Nevertheless, test scores are the best single predictor of an individual's years of education.

In contemporary American society, the amount of schooling that adults complete is also somewhat predictive of their social status. Occupations considered high in prestige (e.g., law, medicine, even corporate business) usually require at least a college degree—16 or more years of education—as a condition of entry. It is partly because intelligence test scores predict years of education so well that they also predict occupational status—and, to a smaller extent, even income (Herrnstein & Murray, 1994; Jencks, 1979). Moreover, many occupations can only be entered through professional schools which base their admissions at least partly on test scores: the MCAT, the GMAT, the LSAT, etc. Individual scores on admission-related tests such as these are certainly correlated with scores on tests of intelligence.

Social status and income. How well do IQ scores (which can be obtained before individuals enter the labor force) predict such outcome measures as the social status or income of adults? This question is complex, in part because another variable also predicts such outcomes: namely, the socioeconomic status (SES) of one's parents. Unsurprisingly, children of privileged families are more likely to attain high social status than those whose parents are poor and less educated. These two predictors (IQ and parental SES) are by no means independent of one another; the correlation between them is around .33 (White, 1982).

One way to look at these relationships is to begin with SES. According to Jencks (1979), measures of parental SES predict about one-third of the variance in young adults' social status and about one-fifth of the variance in their income. About half of this predictive effectiveness depends on the fact that the SES of parents also predicts children's intelligence test scores, which have their own predictive value for social outcomes; the other half comes about in other ways.

We can also begin with IQ scores, which by themselves account for about one-fourth of the social status variance and one-sixth of the income variance. Statistical controls for parental SES eliminate only about a quarter of this predictive power. One way to conceptualize this effect is by comparing the occupational status (or income) of adult brothers who grew up in the same family and hence have the same parental SES. In such cases, the brother with the higher adolescent IQ score is likely to have the higher adult social status and income (Jencks, 1979). This effect, in turn, is substantially mediated by education: the brother with the higher test scores is likely to get more schooling, and hence to be better credentialled as he enters the workplace.

Do these data imply that psychometric intelligence is a major determinant of social status or income? That depends on what one means by "major." In fact, individuals who have the same test scores may differ widely in occupational status and even more widely in income. Consider for a moment the distribution of occupational status scores for all individuals in a population, and then consider the conditional distribution of such scores for just those individuals who test at some given IQ. Jencks (1979) notes that the standard deviation of the latter distribution may still be quite large; in some cases it amounts to about 88% of the standard deviation for the entire population. Viewed from this perspective, psychometric intelligence appears as only one of a great many factors that influence social outcomes.

Job performance. Scores on intelligence tests predict various measures of job performance: supervisor ratings, work samples, etc. Such correlations, which typically lie between $r = .30$ and $r = .50$, are partly restricted by the limited reliability of those measures themselves. They become higher when r is statistically corrected for this unreliability: in one survey of relevant studies (Hunter, 1983), the mean of the corrected correlations was .54. This implies that, across a wide range of occupations, intelligence test performance accounts for some 29% of the variance in job performance.

Although these correlations can sometimes be modified by changing methods of training or aspects of the job itself, intelligence test scores are at least weakly related to job performance in most settings. Sometimes IQ scores are described as the "best available predictor" of that performance. It is worth noting, however, that such tests predict considerably less than half the variance of job-related measures. Other individual characteristics—interpersonal skills, aspects of personality, etc.—are probably of equal or greater importance, but at this point we do not have equally reliable instruments to measure them.

Social outcomes. Psychometric intelligence is negatively correlated with certain socially undesirable outcomes. For example, children with high test scores are less likely than lower-scoring children to engage in juvenile crime. In one study, Moffitt, Gabrielli, Mednick, and Schulsinger (1981) found a correlation of −.19 between IQ scores and number of juvenile offenses in a large Danish sample; with social class controlled, the correlation dropped to −.17. The correlations for most "negative outcome" variables are typically smaller than .20, which means that test scores are associated with less than 4% of their total variance. It is important to realize that the causal links between psychometric ability and social outcomes may be indirect. Children who are unsuccessful in—and hence alienated from—school may be more likely to engage in delinquent behaviors for that very reason, compared to other children who enjoy school and are doing well.

In summary, intelligence test scores predict a wide range of social outcomes with varying degrees of success. Correlations are highest for school achievement, where they account for about a quarter of the variance. They are somewhat lower for job performance, and very low for negatively valued outcomes such as criminality. In general, intelligence tests measure only some of the many personal characteristics that are relevant to life in contemporary America. Those characteristics are never the only influence on outcomes, though in the case of school performance they may well be the strongest.

Test Scores and Measures of Processing Speed

Many recent studies show that the speeds with which people perform very simple perceptual and cognitive tasks are correlated with psychometric intelligence (for reviews see Ceci, 1990; Deary, 1995; Vernon, 1987). In general, people with higher intelligence test scores tend to apprehend, scan, retrieve, and respond to stimuli more quickly than those who score lower.

Cognitive correlates. The modern study of these relations began in the 1970s, as part of the general growth of interest in response time and other chronometric measures of cognition. Many of the new cognitive paradigms required subjects to make same/different judgments or other speeded responses to visual displays. Although those paradigms had not been devised with individual differences in mind, they could be interpreted as providing measures of the speed of certain information processes. Those speeds turned out to correlate with psychometrically-measured verbal ability (Hunt, 1978; Jackson & McClelland, 1979). In some problem solving tasks, it was possible to analyze the subjects' overall response times into theoretically motivated "cognitive components" (Sternberg, 1977); component times could then be correlated with test scores in their own right.

Although the size of these correlations is modest (seldom accounting for more than 10% of the variance), they do increase as the basic tasks were made more complex by requiring increased memory or attentional capacity. For instance, the correlation between paired associate learning and intelligence increases as the pairs are presented at faster rates (Christal, Tirre, & Kyllonen, 1984).

Choice reaction time. In another popular cognitive paradigm, the subject simply moves his or her finger from a "home" button to one of eight other buttons arranged in a semicircle around it; these are marked by small lights that indicate which one is the target on a given trial (Jensen, 1987). Various aspects of the choice reaction times obtained in this paradigm are correlated with scores on intelligence tests, sometimes with values of r as high as −.30 or −.40 (r is negative because higher test scores go with shorter times). Nevertheless, it has proved difficult to make theoretical sense of the overall pattern of correlations, and the results are still hard to interpret (cf. Brody, 1992; Longstreth, 1984).

Somewhat stronger results have been obtained in a variant of Jensen's paradigm devised by Frearson and Eysenck (1986). In this "odd-man-out" procedure, three of the eight lights are illuminated on each trial. Two of these are relatively close to each other while the third is more distant; the subject must press the button corresponding to the more isolated stimulus. Response times in this task show higher correlations with IQ scores than those in Jensen's original procedure, perhaps because it requires more complex forms of spatial judgment.

Inspection time. Another paradigm for measuring processing speed, devised to be relatively independent of response factors, is the method of "inspection time" (IT). In the standard version of this paradigm (Nettelbeck, 1987; Vickers, Nettelbeck & Wilson, 1972), two vertical lines are shown very briefly on each trial, followed by a pattern mask; the subject must judge which line was shorter. For a given subject, IT is defined as the minimum exposure duration (up to the onset of the mask) for which the lines must be displayed if he or she is to meet a pre-established criterion of accuracy—e.g., nine correct trials out of ten.

Inspection times defined in this way are consistently correlated with measures of psychometric intelligence. In a recent meta-analysis, Kranzler and Jensen (1989) reported an overall correlation of −.30 between IQ scores and IT; this rose to −.55 when corrected for measurement error and attenuation. More recent findings confirm this general result (e.g., Bates & Eysenck, 1993; Deary, 1993). IT usually correlates best with performance subtests of intelligence; its correlation with verbal intelligence is usually weaker and sometimes zero.

One apparent advantage of IT over other chronometric methods is that the task itself seems particularly simple. At first glance, it is hard to imagine that any differences in response strategies or stimulus familiarity could affect the outcome. Nevertheless, it seems that they do. Brian Mackenzie and his colleagues (e.g., Mackenzie, Molloy, Martin, Lovegrove, & McNicol, 1991) discovered that some subjects use apparent-movement cues in the basic IT task while others do not; only in the latter group is IT correlated with intelligence test scores. Moreover, standard IT paradigms require an essentially spatial judgment; it is not surprising, then, that they correlate with intelligence tests which emphasize spatial ability. With this in mind, Mackenzie et al. (1991) devised a *verbal* inspection time task based on Posner's classical same-letter/different-letter paradigm (Posner, Boies, Eichelman, & Taylor, 1969). As predicted, the resulting ITs correlated with verbal but not with spatial intelligence. It is clear that the apparently simple IT task actually involves complex modes of information processing (cf. Chaiken, 1993) that are as yet poorly understood.

Neurological measures. Recent research has begun to explore what seem to be still more direct indices of neural processing. Reed and Jensen (1992) have used measures based on visual evoked potentials (VEP) to assess what they call "nerve conduction velocity" (NCV). To estimate that velocity, distance is divided by time: each subject's head length (a rough measure of the distance from the eye to the primary visual cortex) is divided by the latency of an early component (N70 or P100) of his or her evoked potential pattern. In a study with 147 college-student subjects, these NCVs correlated $r = .26$ with scores on an unspeeded test of intelligence. (A statistical correction for the restricted range of subjects raised the correlation to .37.) Other researchers have also reported correlations between VEP parameters and intelligence test scores (Caryl, 1994). Interestingly, however, Reed and Jensen (1993) reported that their estimates of "nerve conduction velocity" were *not* correlated with the same subjects' choice reaction times. Thus while we do not yet understand the basis of the correlation between NCV and psychometric intelligence, it is apparently not just a matter of overall speed.

Problems of interpretation. Some researchers believe that psychometric intelligence, especially *g*, depends directly on what may be called the "neural efficiency" of the brain (Eysenck, 1986; Vernon, 1987). They regard the observed correlations between test scores and measures of processing speed as evidence for their view. If choice reaction times, inspection times, and VEP latencies actually do reflect the speed of basic neural processes, such correlations are only to be expected. In fact, however, the observed patterns of correlation are rarely as simple as this hypothesis would predict. Moreover, it is quite possible that high- and low-IQ individuals differ in other ways that affect speeded performance (cf. Ceci, 1990). Those variables include motivation, response criteria (emphasis on speed vs. accuracy), perceptual strategies (cf. Mackenzie et al., 1991), attentional strategies, and—in some cases—differential familiarity with the material itself. Finally, we do not yet know the direction of causation that underlies such correlations. Do high levels of "neural efficiency" promote the development of intelligence, or do more intelligent people simply find faster ways to carry out perceptual tasks? Or both? These questions are still open.

3. The Genes and Intelligence

In this section of the report we first discuss individual differences generally, without reference to any particular trait. We then focus on intelligence, as measured by conventional IQ tests or other tests intended to measure general cognitive ability. The different and more controversial topic of group differences will be considered in Section 5.

We focus here on the relative contributions of genes and environments to individual differences in particular traits. To avoid misunderstanding, it must be emphasized from the outset that gene action always involves an environment—at least a biochemical environment, and often an ecological one. (For humans, that ecology is usually interpersonal or cultural.) Thus all genetic effects on the development of observable traits are potentially modifiable by environmental input, though the practicability of making such modifications may be another matter. Conversely, all environmental effects on trait development involve the genes or structures to which the genes have contributed. Thus there is always a genetic aspect to the effects of the environment (cf. Plomin & Bergeman, 1991).

Sources of Individual Differences

Partitioning the variation. Individuals differ from one another on a wide variety of traits: familiar examples include height, intelligence, and aspects of personality. Those differences are often of considerable social importance. Many interesting questions can be asked about their nature and origins. One such question is the extent to which they reflect differences among the genes of the individuals involved, as distinguished from differences among the environments to which those individuals have been exposed. The issue here is not whether genes and environments are both essential for the development of a given trait (this is always the case), and it is not about the genes or environment of any particular person. We are concerned only with the observed variation of the trait across individuals in a given population. A figure called the "heritability" (h^2) of the trait represents the proportion of that variation that is associated with genetic differences among the individuals. The remaining variation ($1 - h^2$) is associated with environmental differences and with errors of measurement. These proportions can be estimated by various methods described below.

Sometimes special interest attaches to those aspects of environments that family members have in common (for example, characteristics of the home). The part of the variation that derives from this source, called "shared" variation or c^2, can also be estimated. Still more refined estimates can be made: c^2 is sometimes subdivided into several kinds of shared variation; h^2 is sometimes subdivided into so-called "additive" and "nonadditive" portions (the part that is transmissible from parent to child vs. the part expressed anew in each generation by a unique patterning of genes.) Variation associated with correlations and statistical interactions between genes and environments may also be identifiable. In theory, any of the above estimates may vary with the age of the individuals involved.

A high heritability does not mean that the environment has no impact on the development of a trait, or that learning is not involved. Vocabulary size, for example, is very substantially heritable (and highly correlated with general psychometric intelligence) although every word in an individual's vocabulary is learned. In a society in which plenty of words are available in everyone's environment—especially for individuals who are motivated to seek them out—the number of words that individuals actually learn depends to a considerable extent on their genetic predispositions.

Behavior geneticists have often emphasized the fact that individuals can be active in creating or selecting their own environments. Some describe this process as active or reactive genotype–environment correlation (Plomin, DeFries, & Loehlin, 1977). (The distinction is between the action of the organism in selecting its own environment and the reaction of others to its gene-based traits.) Others suggest that these forms of gene–environment relationship are typical of the way that genes are normally expressed, and simply include them as part of the genetic effect (Roberts, 1967). This is a matter of terminological preference, not a dispute about facts.

How genetic estimates are made. Estimates of the magnitudes of these sources of individual differences are made by exploiting natural and social "experiments" that combine genotypes and environments in informative ways. Monozygotic (MZ) and dizygotic (DZ) twins, for example, can be regarded as experiments of nature. MZ twins are paired individuals of the same age growing up in the same family who have all their genes in common; DZ twins are otherwise similar pairs who have only half their genes in common. Adoptions, in contrast, are experiments of society. They allow one to compare genetically unrelated persons who are growing up in the same family as well as genetically related persons who are growing up in different families. They can also provide information about genotype–environment correlations: in ordinary families genes and environments are correlated because the same parents provide both, whereas in adoptive families one set of parents provides the genes and another the environment. An experiment involving both nature and society is the study of monozygotic twins who have been reared apart (Bouchard, Lykken, McGue, Segal, & Tellegen, 1990; Pedersen, Plomin, Nesselroade, & McClearn, 1992). Relationships in the families of monozygotic twins also offer unique possibilities for analysis (e.g., R. J. Rose, Harris, Christian, & Nance, 1979). Because these comparisons are subject to different sources of potential error, the results of studies involving several kinds of kinship are often analyzed together to arrive at robust overall conclusions. (For general discussions of behavior genetic methods, see Plomin, DeFries, & McClearn, 1990, or Hay, 1985.)

Results for IQ Scores

Parameter estimates. Across the ordinary range of environments in modern Western societies, a sizable part of the variation in intelligence test scores is associated with genetic differences among individuals. Quantitative estimates vary from one study to another, because many are based on small or selective samples. If one simply combines all available correlations in a single analysis, the heritability (h^2) works out to about .50 and the between-family variance (c^2) to about .25 (e.g., Chipuer, Rovine, & Plomin, 1990; Loehlin, 1989). These overall figures are misleading, however, because most of the relevant studies have been done with children. We now know that the heritability of IQ changes with age: h^2 goes up and c^2 goes down from infancy to adulthood (McCartney, Harris, & Bernieri, 1990; McGue, Bouchard, Iacono, & Lykken, 1993). In childhood h^2 and c^2 for IQ are of the order of .45 and .35; by late adolescence h^2 is around .75 and c^2 is quite low (zero in some studies). Substantial environmental variance remains, but it primarily reflects within-family rather than between-family differences.

These adult parameter estimates are based on a number of independent studies. The correlation between MZ twins reared apart, which directly estimates h2, ranged from .68 to .78 in five studies involving adult samples from Europe and the United States (McGue et al., 1993). The correlation between unrelated children reared together in adoptive families, which directly estimates c^2, was approximately zero for adolescents in two adoption studies (Loehlin, Horn, & Willerman, 1989; Scarr & Weinberg, 1978) and .19 in a third (the Minnesota transracial adoption study: Scarr, Weinberg, & Waldman, 1993).

These particular estimates derive from samples in which the lowest socioeconomic levels were under-represented (i.e., there were few very poor families), so the range of between-family differences was smaller than in the population as a whole. This means that we should be cautious in generalizing the findings for between-family effects across the entire social spectrum. The samples were also mostly White, but available data suggest that twin and sibling correlations in African American and similarly selected White samples are more often comparable than not (Loehlin, Lindzey, & Spuhler, 1975).

Why should individual differences in intelligence (as measured by test scores) reflect genetic differences more strongly in adults than they do in children? One possibility is that as individuals grow older their transactions with their environments are increasingly influenced by the characteristics that they bring to those environments themselves, decreasingly by the conditions imposed by family life and social origins. Older persons are in a better position to select their own effective environments, a form of genotype-environment correlation. In any case the popular view that genetic influences on the development of a trait are essentially frozen at conception while the effects of the early environment cumulate inexorably is quite misleading, at least for the trait of psychometric intelligence.

Implications. Estimates of h^2 and c^2 for IQ (or any other trait) are descriptive statistics for the populations studied. (In this respect they are like means and standard deviations.) They are outcome measures, summarizing the results of a great many diverse, intricate, individually variable events and processes, but they can nevertheless be quite useful. They can tell us how much of the variation in a given trait the genes and family environments explain, and changes in them place some constraints on theories of how this occurs. On the other hand they have little to say about specific mechanisms, i.e., about how genetic and environmental differences get translated into individual physiological and psychological differences. Many psychologists and neuroscientists are actively studying such processes; data on heritabilities may give them ideas about what to look for and where or when to look for it.

A common error is to assume that because something is heritable it is necessarily unchangeable. This is wrong. Heritability does not imply immutability. As previously noted, heritable traits can depend on learning, and they may be subject to other environmental effects as well. The value of h^2 can change if the distribution of environments (or genes) in the population is substantially altered. On the other hand, there can be effective environmental changes that do not change heritability at all. If the environment relevant to a given trait improves in a way that affects all members of the population equally, the mean value of the trait will rise without any change in its heritability (because the differences among individuals in the population will stay the same). This has evidently happened for height: the heritability of stature is high, but average heights continue to increase (Olivier, 1980). Something of the sort may also be taking place for IQ scores—the so-called "Flynn effect" discussed in Section 4.

In theory, different subgroups of a population might have different distributions of environments or genes and hence different values of h^2. This seems not to be the case for high and low IQ levels, for which adult heritabilities appear to be much the same (Saudino, Plomin, Pedersen, & McClearn, 1994). It is also possible that an impoverished or suppressive environment could fail to support the development of a trait, and hence restrict individual variation. This could affect estimates of h^2, c^2, or both, depending on the details of the process. Again (as in the case of whole populations), an environmental factor that affected every member of a subgroup equally might alter the group's mean without affecting heritabilities at all.

Where the heritability of IQ is concerned, it has sometimes seemed as if the findings based on differences between group means were in contradiction with those based on correlations. For example, children adopted in infancy into advantaged families tend to have higher IQs in childhood than would have been expected if they had been reared by their birth mothers; this is a mean difference implicating the environment. Yet at the same time their individual resemblance to their birth mothers persists, and this correlation is most plausibly interpreted in genetic terms. There is no real contradiction: the two findings simply call attention to different aspects of the same phenomenon. A sensible account must include both aspects: there is only a single developmental process, and it occurs in individuals. By looking at means or correlations one learns somewhat different but compatible things about the genetic and environmental contributions to that process (Turkheimer, 1991).

As far as behavior genetic methods are concerned, there is nothing unique about psychometric intelligence relative to other traits or abilities. Any reliably measured trait can be analyzed by these methods, and many traits including personality and attitudes have been. The methods are neutral with regard to genetic and environmental sources of variance: if individual differences on a trait are entirely due to environmental factors, the analysis will reveal this. These methods have shown that genes contribute substantially to individual differences in intelligence test performance, and that their role seems to increase from infancy to adulthood. They have also shown that variations in the unique environments of individuals are important, and that between-family variation contributes significantly to observed differences in IQ scores in childhood although this effect diminishes later on. All these conclusions are wholly consistent with the notion that both genes and environment, in complex interplay, are essential to the development of intellectual competence.

4. Environmental Effects on Intelligence

The "environment" includes a wide range of influences on intelligence. Some of those variables affect whole populations, while others contribute to individual differences within a given group. Some of them are social, some are biological; at this point some are still mysterious. It may also happen that the proper interpretation of an environmental variable requires the simultaneous consideration of genetic effects. Nevertheless, a good deal of solid information is available.

Social Variables

It is obvious that the cultural environment—how people live, what they value, what they do—has a significant effect on the intellectual skills developed by individuals. Rice farmers in Liberia are good at estimating quantities of rice (Gay & Cole, 1967); children in Botswana, accustomed to story-telling, have excellent memories for stories

(Dube, 1982). Both these groups were far ahead of American controls on the tasks in question. On the other hand Americans and other Westernized groups typically outperform members of traditional societies on psychometric tests, even those designed to be "culture-fair."

Cultures typically differ from one another in so many ways that particular differences can rarely be ascribed to single causes. Even comparisons between subpopulations can be difficult to interpret. If we find that middle-class and poor Americans differ in their scores on intelligence tests, it is easy to suppose that the environmental difference has caused the IQ difference (i.e., that growing up in the middle class produces higher psychometric intelligence than growing up poor). But there may also be an opposite direction of causation: individuals can come to be in one environment or another because of differences in their own abilities. Waller (1971) has shown, for example, that adult sons whose IQ scores are above those of their fathers tend to have higher social-class status than those fathers; conversely, sons with IQ scores below their fathers' tend to have lower social-class status. Since all the subjects grew up with their fathers, the IQ differences in this study cannot have resulted from class-related differences in childhood experience. Rather, those differences (or other factors correlated with them) seem to have had an influence on the status that they achieved. Such a result is not surprising, given the relation between test scores and years of education reviewed in Section 2.

Occupation. In Section 2 we noted that intelligence test scores predict occupational level, not only because some occupations require more intelligence than others but also because admission to many professions depends on test scores in the first place. There can also be an effect in the opposite direction, i.e., workplaces may affect the intelligence of those who work in them. Kohn and Schooler (1973), who interviewed some 3,000 men in various occupations (farmers, managers, machinists, porters, etc.), argued that more "complex" jobs produce more "intellectual flexibility" in the individuals who hold them. Although the issue of direction of effects was not fully resolved in their study—and perhaps not even in its longitudinal follow-up (Kohn & Schooler, 1983)—this remains a plausible suggestion.

Among other things, Kohn and Schooler's hypothesis may help us understand urban/rural differences. A generation ago these were substantial in the United States, averaging about 6 IQ points or 0.4 standard deviations (Terman & Merrill, 1937; Seashore, Wesman, & Doppelt, 1950). In recent years the difference has declined to about 2 points (Kaufman & Doppelt, 1976; Reynolds, Chastain, Kaufman, & McLean, 1987). In all likelihood this urban/rural convergence primarily reflects environmental changes: a decrease in rural isolation (due to increased travel and mass communications), an improvement in rural schools, the greater use of technology on farms. All these changes can be regarded as increasing the "complexity" of the rural environment in general or of farm work in particular. (However, processes with a genetic component—e.g., changes in the selectivity of migration from farm to city—cannot be completely excluded as contributing factors.)

Schooling. Attendance at school is both a dependent and an independent variable in relation to intelligence. On the one hand, children with higher test scores are less likely to drop out and more likely to be promoted from grade to grade and then to attend college. Thus the number of years of education that adults complete is roughly predictable from their childhood scores on intelligence tests. On the other hand, schooling itself changes mental abilities, including those abilities measured on psychometric tests. This is obvious for tests like the SAT that are explicitly designed to assess school learning, but it is almost equally true of intelligence tests themselves.

The evidence for the effect of schooling on intelligence test scores takes many forms (Ceci, 1991). When children of nearly the same age go through school a year apart (because of birthday-related admission criteria), those who have been in school longer have higher mean scores. Children who attend school intermittently score below those who go regularly, and test performance tends to drop over the summer vacation. A striking demonstration of this effect appeared when the schools in one Virginia county closed for several years in the 1960s to avoid integration, leaving most Black children with no formal education at all. Compared to controls, the intelligence-test scores of these children dropped by about 0.4 standard deviations (6 points) per missed year of school (Green, Hoffman, Morse, Hayes, & Morgan, 1964).

Schools affect intelligence in several ways, most obviously by transmitting information. The answers to questions like "Who wrote Hamlet?" and "What is the boiling point of water?" are typically learned in school, where some pupils learn them more easily and thoroughly than others. Perhaps at least as important are certain general skills and attitudes: systematic problem-solving, abstract thinking, categorization, sustained attention to material of little intrinsic interest, and repeated manipulation of basic symbols and operations. There is no doubt that schools promote and permit the development of significant intellectual skills, which develop to different extents in different children. It is because tests of intelligence draw on many of those same skills that they predict school achievement as well as they do.

To achieve these results, the school experience must meet at least some minimum standard of quality. In very poor schools, children may learn so little that they fall farther behind the national IQ norms for every year of attendance. When this happens, older siblings have systematically lower scores than their younger counterparts. This pattern of scores appeared in at least one rural Georgia school system in the 1970s (Jensen, 1977). Before desegregation, it must have been characteristic of many of the schools attended by Black pupils in the South. In a study based on Black children who had moved to Philadelphia at various ages during this period, Lee (1951) found that their IQ scores went up more than half a point for each year that they were enrolled in the Philadelphia system.

Interventions. Intelligence test scores reflect a child's standing relative to others in his or her age cohort. Very poor or interrupted schooling can lower that standing substantially; are there also ways to raise it? In fact many interventions have been shown to raise test scores and mental ability "in the short run" (i.e., while the program itself was in progress), but long-run gains have proved more elusive. One noteworthy example of (at least short-run) success was the Venezuelan Intelligence Project (Herrnstein, Nickerson, de Sanchez, & Swets, 1986), in which hundreds of seventh-grade children from under-privileged backgrounds in that country were exposed to an extensive, theoretically-based curriculum focused on thinking skills. The intervention produced substantial gains on a wide range of tests, but there has been no follow-up.

Children who participate in "Head Start" and similar programs are exposed to various school-related materials and experiences for one or two years. Their test scores often go up during the course of the program, but these gains fade with time. By the end of elementary school, there are usually no significant IQ or achievement-test differences between children who have been in such programs and controls who have not. There may, however, be other differences. Follow-up studies suggest that children who participated in such programs as preschoolers are less likely to be assigned to special education, less likely to be held back in grade, and more likely to finish high school than matched controls (Consortium for Longitudinal Studies, 1983; Darlington, 1986; but see Locurto, 1991).

More extensive interventions might be expected to produce larger and more lasting effects, but few such programs have been evaluated systematically. One of the more successful is the Carolina Abecedarian Project (Campbell & Ramey, 1994), which provided a group of children with enriched environments from early infancy through preschool and also maintained appropriate controls. The test scores of the enrichment-group children were already higher than those of controls at age two; they were still some 5 points higher at age 12, seven years after the end of the intervention. Importantly, the enrichment group also outperformed the controls in academic achievement.

Family environment. No one doubts that normal child development requires a certain minimum level of responsible care. Severely deprived, neglectful, or abusive environments must have negative effects on a great many aspects—including intellectual aspects—of development. Beyond that minimum, however, the role of family experience is now in serious dispute (Baumrind, 1993; Jackson, 1993; Scarr, 1992, 1993). Psychometric intelligence is a case in point. Do differences between children's family environments (within the normal range) produce differences in their intelligence test performance? The problem here is to disentangle causation from correlation. There is no doubt that such variables as resources of the home (Gottfried, 1984) and parents' use of language (Hart & Risley, 1992, in press) are correlated with children's IQ scores, but such correlations may be mediated by genetic as well as (or instead of) environmental factors.

Behavior geneticists frame such issues in quantitative terms. As noted in Section 3, environmental factors certainly contribute to the overall variance of psychometric intelligence. But how much of that variance results from differences between families, as contrasted with the varying experiences of different children in the same family? Between-family differences create what is called "shared variance" or c^2 (all children in a family share the same home and the same parents). Recent twin and adoption studies suggest that while the value of c^2 (for IQ scores) is substantial in early childhood, it becomes quite small by late adolescence.

These findings suggest that differences in the life styles of families—whatever their importance may be for many aspects of children's lives—make little long-term difference for the skills measured by intelligence tests. We should note, however, that low-income and non-White families are poorly represented in existing adoption studies as well as in most twin samples. Thus it is not yet clear whether these surprisingly small values of (adolescent) c^2 apply to the population as a whole. It remains possible that, across the full range of income and ethnicity, between-family differences have more lasting consequences for psychometric intelligence.

Biological Variables

Every individual has a biological as well as a social environment, one that begins in the womb and extends throughout life. Many aspects of that environment can affect intellectual development. We now know that a number of biological factors—malnutrition, exposure to toxic substances, various prenatal and perinatal stressors—result in lowered psychometric intelligence under at least some conditions.

Nutrition. There has been only one major study of the effects of prenatal malnutrition (i.e., malnutrition of the mother during pregnancy) on long-term intellectual development. Stein, Susser, Saenger, and Marolla (1975) analyzed the test scores of Dutch 19-year-old males in relation to a wartime famine that had occurred in the winter of 1944–45, just before their birth. In this very large sample (made possible by a universal military induction requirement), exposure to the famine had no effect on adult intelligence. Note, however, that the famine itself lasted only a few months; the subjects were exposed to it prenatally but not after birth.

In contrast, prolonged malnutrition during childhood does have long-term intellectual effects. These have not been easy to establish, in part because many other unfavorable socioeconomic conditions are often associated with chronic malnutrition (Ricciuti, 1993; but cf. Sigman, 1995). In one intervention study, however, preschoolers in two Guatemalan villages (where undernourishment is common) were given ad lib access to a protein dietary supplement for several years. A decade later, many of these children (namely, those from the poorest socioeconomic levels) scored significantly higher on school-related achievement tests than comparable controls (Pollitt, Gorman, Engle, Martorell, & Rivera, 1993). It is worth noting that the effects of poor nutrition on intelligence may well be indirect. Malnourished children are typically less responsive to adults, less motivated to learn, and less active in exploration than their more adequately nourished counterparts.

Although the degree of malnutrition prevalent in these villages rarely occurs in the United States, there may still be nutritional influences on intelligence. In studies of so-called "micro-nutrients," experimental groups of children have been given vitamin/mineral supplements while controls got placebos. In many of these studies (e.g., Schoenthaler, Amos, Eysenck, Peritz, & Yudkin, 1991), the experimental children showed test-score gains that significantly exceeded the controls. In a somewhat different design, Rush, Stein, Susser, and Brody (1980) gave dietary supplements of liquid protein to pregnant women who were thought to be at risk for delivering low birth-weight babies. At one year of age, the babies born to these mothers showed faster habituation to visual patterns than did control infants. (Other research has shown that infant habituation rates are positively correlated with later psychometric test scores: Columbo, 1993.) Although these results are encouraging, there has been no long-term follow-up of such gains.

Lead. Certain toxins have well-established negative effects on intelligence. Exposure to lead is one such factor. In one long-term study (Baghurst et al., 1992; McMichael et al., 1988), the blood lead levels of children growing up near a lead smelting plant were substantially and negatively correlated with intelligence test scores throughout childhood. No "threshold dose" for the effect of lead appears in such studies. Although ambient lead levels in the United States have been reduced in recent years, there is reason to believe that some American children—especially those in inner cities—may still be at risk from this source (cf. Needleman, Geiger, & Frank, 1985).

Alcohol. Extensive prenatal exposure to alcohol (which occurs if the mother drinks heavily during pregnancy) can give rise to fetal alcohol syndrome, which includes mental retardation as well as a range of physical symptoms. Smaller "doses" of prenatal alcohol may have negative effects on intelligence even when the full syndrome does not appear. Streissguth, Barr, Sampson, Darby, and Martin (1989) found that mothers who reported consuming more than 1.5 oz. of alcohol daily during pregnancy had children who scored some 5 points below controls at age four. Prenatal exposure to aspirin and antibiotics had similar negative effects in this study.

Perinatal factors. Complications at delivery and other negative perinatal factors may have serious consequences for development. Nevertheless, because they occur only rarely, they contribute relatively little to the population variance of intelligence (Broman, Nichols, & Kennedy, 1975). Down's syndrome, a chromosomal abnormality that produces serious mental retardation, is also rare enough to have little impact on the overall distribution of test scores.

The correlation between birth weight and later intelligence deserves particular discussion. In some cases low birth weight simply reflects premature delivery; in others, the infant's size is below normal for its gestational age. Both factors apparently contribute to the tendency of low-birth-weight infants to have lower test scores in later childhood (Lubchenko, 1976). These correlations are small, ranging from .05 to .13 in different groups (Broman et al., 1975). The effects of low birth weight are substantial only when it is very low indeed (less than 1,500 gm). Premature babies born at these very low birth weights are behind controls on most developmental measures; they often have severe or permanent intellectual deficits (Rosetti, 1986).

Continuously Rising Test Scores

Perhaps the most striking of all environmental effects is the steady worldwide rise in intelligence test performance. Although many psychometricians had noted these gains, it was James Flynn (1984, 1987) who first described them systematically. His analysis shows that performance has been going up ever since testing began. The "Flynn effect" is now very well documented, not only in the United States but in many other technologically advanced countries. The average gain is about 3 IQ points per decade—more than a full standard deviation since, say, 1940.

Although it is simplest to describe the gains as increases in population IQ, this is not exactly what happens. Most intelligence tests are "restandardized" from time to time, in part to keep up with these very gains. As part of this process the mean score of the new standardization sample is typically set to 100 again, so the increase more or less disappears from view. In this context, the Flynn effect means that if 20 years have passed since the last time the test was standardized, people who now score 100 on the new version would probably average about 106 on the old one.

The sheer extent of these increases is remarkable, and the rate of gain may even be increasing. The scores of 19-year-olds in the Netherlands, for example, went up more than 8 points—over half a standard deviation—between 1972 and 1982. What's more, the largest gains appear on the types of tests that were specifically designed to be free of cultural influence (Flynn, 1987). One of these is Raven's Progressive Matrices, an untimed nonverbal test that many psychometricians regard as a good measure of *g*.

These steady gains in intelligence test performance have not always been accompanied by corresponding gains in school achievement. Indeed, the relation between intelligence and achievement test scores can be complex. This is especially true for the Scholastic Aptitude Test (SAT), in part because the ability range of the students who take the SAT has broadened over time. That change explains some portion—not all—of the prolonged decline in SAT scores that took place from the mid-1960s to the early 1980s, even as IQ scores were continuing to rise (Flynn, 1984). Meanwhile, however, other more representative measures show that school achievement levels have held steady or in some cases actually increased (Herrnstein & Murray, 1994). The National Assessment of Educational Progress (NAEP), for example, shows that the average reading and math achievement of American 13- and 17-year-olds improved somewhat from the early 1970s to 1990 (Grissmer, Kirby, Berends, & Williamson, 1994). An analysis of these data by ethnic group, reported in Section 5, shows that this small overall increase actually reflects very substantial gains by Blacks and Latinos combined with little or no gain by Whites.

The consistent IQ gains documented by Flynn seem much too large to result from simple increases in test sophistication. Their cause is presently unknown, but three interpretations deserve our consideration. Perhaps the most plausible of these is based on the striking cultural differences between successive generations. Daily life and occupational experience both seem more "complex" (Kohn & Schooler, 1973) today than in the time of our parents and grandparents. The population is increasingly urbanized; television exposes us to more information and more perspectives on more topics than ever before; children stay in school longer; and almost everyone seems to be encountering new forms of experience. These changes in the complexity of life may have produced corresponding changes in complexity of mind, and hence in certain psychometric abilities.

A different hypothesis attributes the gains to modern improvements in nutrition. Lynn (1990) points out that large nutritionally-based increases in height have occurred during the same period as the IQ gains: perhaps there have been increases in brain size as well. As we have seen, however, the effects of nutrition on intelligence are themselves not firmly established.

The third interpretation addresses the very definition of intelligence. Flynn himself believes that real intelligence—whatever it may be—cannot have increased as much as these data would suggest. Consider, for example, the number of individuals who have IQ scores of 140 or more. (This is slightly above the cutoff used by L. M. Terman [1925] in his famous longitudinal study of "genius.") In 1952 only 0.38% of Dutch test takers had IQs over 140; in 1982, scored by the same norms, 9.12% exceeded this figure! Judging by these criteria, the Netherlands should now be experiencing "a cultural renaissance too great to be overlooked" (Flynn, 1987, p. 187). So too should France, Norway, the United States, and many other countries. Because Flynn (1987) finds this conclusion implausible or absurd, he argues that what has risen cannot be intelligence itself but only a minor sort of "abstract problem solving ability." The issue remains unresolved.

Individual Life Experiences

Although the environmental variables that produce large differences in intelligence are not yet well understood, genetic studies assure us that they exist. With a heritability well below 1.00, IQ must be subject to substantial environmental influences. Moreover, available heritability estimates apply only within the range of environments that are well-represented in the present population. We already know that some relatively rare conditions, like those reviewed earlier, have large negative effects on intelligence. Whether there are (now equally rare) conditions that have large positive effects is not known.

As we have seen, there is both a biological and a social environment. For any given child, the social factors include not only an overall cultural/social/school setting and a particular family but also a unique "micro-environment" of experiences that are shared with no one else. The adoption studies reviewed in Section 3 show that family variables—differences in parenting style, in the resources of the home, etc.—have smaller long-term effects than we once supposed. At least among people who share a given SES level and a given culture, it seems to be unique individual experience that makes the largest environmental contribution to adult IQ differences.

We do not yet know what the key features of those micro-environments may be. Are they biological? Social? Chronic? Acute? Is there something especially important in the earliest relations between the infant and its caretakers? Whatever the critical variables may be, do they interact with other aspects of family life? Of culture? At this point we cannot say, but these questions offer a fertile area for further research.

5. Group Differences

Group means have no direct implications for individuals. What matters for the next person you meet (to the extent that test scores matter at all) is that person's own particular score, not the mean of some reference group to which he or she happens to belong. The commitment to evaluate people on their own individual merit is central to a democratic society. It also makes quantitative sense. The distributions of different groups inevitably overlap, with the range of scores within any one group always wider than the mean differences between any two groups. In the case of intelligence test scores, the variance attributable to individual differences far exceeds the variance related to group membership (Jensen, 1980).

Because claims about ethnic differences have often been used to rationalize racial discrimination in the past, all such claims must be subjected to very careful scrutiny. Nevertheless, group differences continue to be the subject of intense interest and debate. There are many reasons for this interest: some are legal and political, some social and psychological. Among other things, facts about group differences may be relevant to the need for (and the effectiveness of) affirmative action programs. But while some recent discussions of intelligence and ethnic differences (e.g., Herrnstein & Murray, 1994) have made specific policy recommendations in this area, we will not do so here. Such recommendations are necessarily based on political as well as scientific considerations, and so fall outside the scope of this report.

Besides European Americans ("Whites"), the ethnic groups to be considered are Chinese and Japanese Americans, Hispanic Americans ("Latinos"), Native Americans ("Indians"), and African Americans ("Blacks"). These groups (we avoid the term "race") are defined and self-defined by social conventions based on ethnic origin as well as on observable physical characteristics such as skin color. None of them are internally homogeneous. Asian Americans, for example, may have roots in many different cultures: not only China and Japan but also Korea, Laos, Vietnam, the Philippines, India, and Pakistan. Hispanic Americans, who share a common linguistic tradition, actually differ along many cultural dimensions. In their own minds they may be less "Latinos" than Puerto Ricans, Mexican Americans, Cuban Americans, or representatives of other Latin cultures. "Native American" is an even more diverse category, including a great many culturally distinct tribes living in a wide range of environments.

Although males and females are not ethnic or cultural groups, possible sex differences in cognitive ability have also been the subject of widespread interest and discussion. For this reason, the evidence relevant to such differences is briefly reviewed in the next section.

Sex Differences

Most standard tests of intelligence have been constructed so that there are no overall score differences between females and males. Some recent studies do report sex differences in IQ, but the direction is variable and the effects are small (Held, Alderton, Foley, & Segall, 1993; Lynn, 1994). This overall equivalence does not imply equal performance on every individual ability. While some tasks show no sex differences, there are others where small differences appear and a few where they are large and consistent.

Spatial and quantitative abilities. Large differences favoring males appear on visual-spatial tasks like mental rotation and spatiotemporal tasks like tracking a moving object through space (Law, Pellegrino, & Hunt, 1993; Linn & Petersen, 1985). The sex difference on mental rotation tasks is substantial: a recent meta-analysis (Masters & Sanders, 1993) puts the effect size at $d = 0.9$. (Effect sizes are measured in standard deviation units. Here, the mean of the male distribution is nearly one standard deviation above that for females.) Males' achievement levels on movement-related and visual-spatial tests are relevant to their generally better performance in tasks that involve aiming and throwing (Jardine & Martin, 1983).

Some quantitative abilities also show consistent differences. Females have a clear advantage on quantitative tasks in the early years of school (Hyde, Fennema, & Lamon, 1990), but this reverses sometime before puberty; males then maintain their superior performance into old age. The math portion of the Scholastic Aptitude Test shows a substantial advantage for males ($d = 0.33$ to 0.50), with many more males scoring in the highest ranges (Benbow, 1988; Halpern, 1992). Males also score consistently higher on tests of proportional and mechanical reasoning (Meehan, 1984; Stanley, Benbow, Brody, Dauber, & Lupkowski, 1992).

Verbal abilities. Some verbal tasks show substantial mean differences favoring females. These include synonym generation and verbal fluency (e.g., naming words that start with a given letter), with effect sizes ranging from $d = 0.5$ to 1.2 (Gordon & Lee, 1986; Hines, 1990). On average females score higher on college achievement tests in literature, English composition, and Spanish (Stanley, 1993); they also excel at reading and spelling. Many more males than females are diagnosed with dyslexia and other reading disabilities (Sutaria, 1985), and there are many more male stutterers (Yairi & Ambrose, 1992). Some memory tasks also show better performance by females, but the size (and perhaps even the direction) of the effect varies with the type of memory being assessed.

Causal factors. There are both social and biological reasons for these differences. At the social level there are both subtle and overt differences between the experiences, expectations, and gender roles of females and males. Relevant environmental differences appear soon after birth. They range from the gender-differentiated toys that children regularly receive to the expectations of adult life with which they are presented, from gender-differentiated household and leisure activities to assumptions about differences in basic ability. Models that include many of these psychosocial variables have been successful in predicting academic achievement (Eccles, 1987).

Many biological variables are also relevant. One focus of current research is on differences in the sizes or shapes of particular neural structures. Numerous sexually dimorphic brain structures have now been identified, and they may well have implications for cognition. There are, for example, sex-related differences in the sizes of some portions of the corpus callosum; these differences are correlated with verbal fluency (Hines, Chiu, McAdams, Bentler, & Lipcamon, 1992). Recent brain imaging studies have found what may be differences in the lateralization of language (Shaywitz et al., 1995). Note that such differences in neural structure could result from differences in patterns of life experience as well as from genetically-driven mechanisms of brain development; moreover, brain development and experience may have bidirectional effects on each other. This research area is still in a largely exploratory phase.

Hormonal influences. The importance of prenatal exposure to sex hormones is well established. Hormones influence not only the developing genitalia but also the brain and certain immune system structures (Geschwind & Galaburda, 1987; Halpern & Cass, 1994). Several studies have tested individuals who were exposed to abnormally high androgen levels in utero, due to a condition known as congenital adrenal hyperplasia (CAH). Adult CAH females score significantly higher than controls on tests of spatial ability (Resnick, Berenbaum, Gottesman & Bouchard, 1986); CAH girls play more with "boys' toys" and less with "girls' toys" than controls (Berenbaum & Hines, 1992).

Other experimental paradigms confirm the relevance of sex hormones for performance levels in certain skills.

Christiansen and Knussman (1987) found testosterone levels in normal males to be correlated positively (about .20) with some measures of spatial ability and negatively (about −.20) with some measures of verbal ability. Older males given testosterone show improved performance on visual-spatial tests (Janowsky, Oviatt, & Orwoll, 1994). Many similar findings have been reported, though the effects are often nonlinear and complex (Gouchie & Kimura, 1991; Nyborg, 1984). It is clear that any adequate model of sex differences in cognition will have to take both biological and psychological variables (and their interactions) into account.

Mean Scores of Different Ethnic Groups

Asian Americans. In the years since the Second World War, Asian Americans—especially those of Chinese and Japanese extraction—have compiled an outstanding record of academic and professional achievement. This record is reflected in school grades, in scores on content-oriented achievement tests like the SAT and GRE, and especially in the disproportionate representation of Asian Americans in many sciences and professions. Although it is often supposed that these achievements reflect correspondingly high intelligence test scores, this is not the case. In more than a dozen studies from the 1960s and 1970s analyzed by Flynn (1991), the mean IQs of Japanese and Chinese American children were always around 97 or 98; none was over 100. Even Lynn (1993), who argues for a slightly higher figure, concedes that the achievements of these Asian Americans far outstrip what might have been expected on the basis of their test scores.

It may be worth noting that the interpretation of test scores obtained by Asians in Asia has been controversial in its own right. Lynn (1982) reported a mean Japanese IQ of 111 while Flynn (1991) estimated it to be between 101 and 105. Stevenson et al. (1985), comparing the intelligence-test performance of children in Japan, Taiwan, and the United States, found no substantive differences at all. Given the general problems of cross-cultural comparison, there is no reason to expect precision or stability in such estimates. Nevertheless, some interest attaches to these particular comparisons: they show that the well-established differences in school achievement among the same three groups (Chinese and Japanese children are much better at math than American children) do not simply reflect differences in psychometric intelligence. Stevenson, Lee, and Stigler (1986) suggest that they result from structural differences in the schools of the three nations as well as from varying cultural attitudes toward learning itself. It is also possible that spatial ability—in which Japanese and Chinese obtain somewhat higher scores than Americans—plays a particular role in the learning of mathematics.

One interesting way to assess the achievements of Chinese and Japanese Americans is to reverse the usual direction of prediction. Data from the 1980 census show that the proportion of Chinese Americans employed in managerial, professional, or technical occupations was 55% and that of Japanese was 46%. (For Whites, the corresponding figure was 34%.) Using the well-established correlation between intelligence test scores and occupational level, Flynn (1991, p. 99) calculated the mean IQ that a hypothetical White group "would have to have" to predict the same proportions of upper-level employment. He found that the occupational success of these Chinese Americans—whose mean IQ was in fact slightly below 100—was what would be expected of a White group with an IQ of almost 120! A similar calculation for Japanese Americans shows that their level of achievement matched that of Whites averaging 110. These "overachievements" serve as sharp reminders of the limitations of IQ-based prediction. Various aspects of Chinese American and Japanese American culture surely contribute to them (Schneider, Hieshima, Lee, & Plank, 1994); gene-based temperamental factors could conceivably be playing a role as well (Freedman & Freedman, 1969).

Hispanic Americans. Hispanic immigrants have come to America from many countries. In 1993, the largest Latino groups in the continental United States were Mexican Americans (64%), Puerto Ricans (11%), Central and South Americans (13%), and Cubans (5%) (U.S. Bureau of the Census, 1994). There are very substantial cultural differences among these nationality groups, as well as differences in academic achievement (Duran, 1983; United States National Commission for Employment Policy, 1982). Taken together, Latinos make up the second largest and the fastest-growing minority group in America (Davis, Haub, & Willette, 1983; Eyde, 1992).

In the United States, the mean intelligence test scores of Hispanics typically lie between those of Blacks and Whites. There are also differences in the patterning of scores across different abilities and subtests (Hennessy & Merrifield, 1978; Lesser, Fifer, & Clark, 1965). Linguistic factors play a particularly important role for Hispanic Americans, who may know relatively little English. (By one estimate, 25% of Puerto Ricans and Mexican Americans and at least 40% of Cubans speak English "not well" or "not at all" [Rodriguez, 1992]). Even those who describe themselves as bilingual may be at a disadvantage if Spanish was their first and best-learned language. It is not surprising that Latino children typically score higher on the performance than on the verbal subtests of the English-based Wechsler Intelligence Scale for Children—Revised (WISC–R; Kaufman, 1994). Nevertheless, the predictive validity of Latino test scores is not negligible. In young children, the WISC–R has reasonably high correlations with school achievement measures (McShane & Cook, 1985). For high school students of moderate to high English proficiency, standard aptitude tests predict first-year college grades about as well as they do for non-Hispanic Whites (Pennock-Roman, 1992).

Native Americans. There are a great many culturally distinct North American Indian tribes (Driver, 1969), speaking some 200 different languages (Leap, 1981). Many Native Americans live on reservations, which themselves represent a great variety of ecological and cultural settings. Many others presently live in metropolitan areas (Brandt, 1984). Although few generalizations can be appropriate across so wide a range, two or three points seem fairly well established. The first is a specific relation between ecology and cognition: the Inuit (Eskimo) and other groups that live in the arctic tend to have particularly high visual-spatial skills. (For a review see McShane & Berry, 1988.) Moroever, there seem to be no substantial sex differences in those skills (Berry, 1974). It seems likely that this represents an adaptation—genetic or learned or both—to the difficult hunting, traveling, and living conditions that characterize the arctic environment.

On the average, Indian children obtain relatively low scores on tests of verbal intelligence, which are often administered in school settings. The result is a performance-test/verbal-test discrepancy similar to that exhibited by Hispanic Americans and other groups whose first language is generally not English. Moreover, many Indian children suffer from chronic middle-ear infection (otitis media), which is "the leading identifiable disease among Indians since record-keeping began in 1962" (McShane & Plas, 1984a, p. 84). Hearing loss can have marked negative effects on verbal test performance (McShane & Plas, 1984b).

African Americans. The relatively low mean of the distribution of African American intelligence test scores has been discussed for many years. Although studies using different tests and samples yield a range of results, the Black mean is typically about one standard deviation (about 15 points) below that of Whites (Jensen, 1980; Loehlin et al., 1975; Reynolds et al., 1987). The difference is largest on those tests (verbal or nonverbal) that best represent the general intelligence factor *g* (Jensen, 1985). It is possible, however, that this differential is diminishing. In the most recent restandardization of the Stanford-Binet test, the Black/White differential was 13 points for younger children and 10 points for older children (Thorndike, Hagen, & Sattler, 1986). In several other studies of children since 1980, the Black mean has consistently been over 90 and the differential has been in single digits (Vincent, 1991). Larger and more definitive studies are needed before this trend can be regarded as established.

Another reason to think the IQ mean might be changing is that the Black/White differential in *achievement* scores has diminished substantially in the last few years. Consider, for example, the mathematics achievement of 17-year-olds as measured by the National Assessment of Educational Progress (NAEP). The differential between Black and White scores, about 1.1 standard deviations as recently as 1978, had shrunk to .65 *SD* by 1990 (Grissmer et al., 1994) because of Black gains. Hispanics showed similar but smaller gains; there was little change in the scores of Whites. Other assessments of school achievement also show substantial recent gains in the performance of minority children.

In their own analysis of these gains, Grissmer et al. (1994) cite both demographic factors and the effects of public policy. They found the level of parents' education to be a particularly good predictor of children's school achievement; that level increased for all groups between 1970 and 1990, but most sharply for Blacks. Family size was another good predictor (children from smaller families tend to achieve higher scores); here too, the largest change over time was among Blacks. Above and beyond these demographic effects, Grissmer et al. believe that some of the gains can be attributed to the many specific programs, geared to the education of minority children, that were implemented during that period.

Test bias. It is often argued that the lower mean scores of African Americans reflect a bias in the intelligence tests themselves. This argument is right in one sense of "bias" but wrong in another. To see the first of these, consider how the term is used in probability theory. When a coin comes up heads consistently for any reason it is said to be "biased," regardless of any consequences that the outcome may or may not have. In this sense the Black/White score differential is *ipso facto* evidence of what may be called "outcome bias." African Americans are subject to outcome bias not only with respect to tests but along many dimensions of American life. They have the short end of nearly every stick: average income, representation in high-level occupations, health and health care, death rate, confrontations with the legal system, and so on. With this situation in mind, some critics regard the test score differential as just another example of a pervasive outcome bias that characterizes our society as a whole (Jackson, 1975; Mercer, 1984). Although there is a sense in which they are right, this critique ignores the particular social purpose that tests are designed to serve.

From an educational point of view, the chief function of mental tests is as *predictors* (Section 2). Intelligence tests predict school performance fairly well, at least in American schools as they are now constituted. Similarly, achievement tests are fairly good predictors of performance in college and postgraduate settings. Considered in this light, the relevant question is whether the tests have a "predictive bias" against Blacks. Such a bias would exist if African American performance on the criterion variables (school achievement, college GPA, etc.) were systematically higher than the same subjects' test scores would predict. This is not the case. The actual regression lines (which show the mean criterion performance for individuals who got various scores on the predictor) for Blacks do not lie above those for Whites; there is even a slight tendency in the other direction (Jensen, 1980; Reynolds & Brown, 1984). Considered as predictors of future performance, the tests do not seem to be biased against African Americans.

Characteristics of tests. It has been suggested that various aspects of the way tests are formulated and administered may put African Americans at a disadvantage. The language of testing is a standard form of English with which some Blacks may not be familiar; specific vocabulary items are often unfamiliar to Black children; the tests are often given by White examiners rather than by more familiar Black teachers; African Americans may not be motivated to work hard on tests that so clearly reflect White values; the time demands of some tests may be alien to Black culture. (Similar suggestions have been made in connection with the test performance of Hispanic Americans, e.g., Rodriguez, 1992.) Many of these suggestions are plausible, and such mechanisms may play a role in particular cases. Controlled studies have shown, however, that none of them contributes substantially to the Black/White differential under discussion here (Jensen, 1980; Reynolds & Brown, 1984; for a different view see Helms, 1992). Moreover, efforts to devise reliable and valid tests that would minimize disadvantages of this kind have been unsuccessful.

Interpreting Group Differences

If group differences in test performance do not result from the simple forms of bias reviewed above, what is responsible for them? The fact is that we do not know. Various explanations have been proposed, but none is generally accepted. It is clear, however, that these differences—whatever their origin—are well within the range of effect sizes that can be produced by environmental factors. The Black/White differential amounts to one standard deviation or less, and we know that environmental factors have recently raised mean test scores in many populations by at least that much (Flynn, 1987: see Section 4). To be sure, the "Flynn effect" is itself poorly understood: it may reflect generational changes in culture, improved nutrition, or other factors as yet unknown. Whatever may be responsible for it, we cannot exclude the possibility that the same factors play a role in contemporary group differences.

Socioeconomic factors. Several specific environmental/cultural explanations of those differences have been proposed. All of them refer to the general life situation in which contemporary African Americans find themselves, but that situation can be described in several different ways. The simplest such hypothesis can be framed in economic terms. On the average, Blacks have lower incomes than Whites; a much higher proportion of them are poor. It is plausible to suppose that many inevitable aspects of poverty—poor nutrition, frequently inadequate prenatal care, lack of intellectual resources—have negative effects on children's developing intelligence. Indeed, the correlation between "socioeconomic status" (SES) and scores on intelligence tests is well-known (White, 1982).

Several considerations suggest that this cannot be the whole explanation. For one thing, the Black/White differential in test scores is not eliminated when groups or individuals are matched for SES (Loehlin et al., 1975). Moreover, the data reviewed in Section 4 suggest that—if we exclude extreme conditions—nutrition and other biological factors that may vary with SES account for relatively little of the variance in such scores. Finally, the (relatively weak) relationship between test scores and income is much more complex than a simple SES hypothesis would suggest. The living conditions of children result in part from the accomplishments of their parents: If the skills measured by psychometric tests actually matter for those accomplishments, intelligence is affecting SES rather than the other way around. We do not know the magnitude of these various effects in various populations, but it is clear that no model in which "SES" directly determines "IQ" will do.

A more fundamental difficulty with explanations based on economics alone appears from a different perspective. To imagine that any simple income- and education-based index can adequately describe the situation of African Americans is to ignore important categories of experience. The sense of belonging to a group with a distinctive culture—one that has long been the target of oppression—and the awareness or anticipation of racial discrimination are profound personal experiences, not just aspects of socioeconomic status. Some of these more deeply rooted differences are addressed by other hypotheses, based on caste and culture.

Caste-like minorities. Most discussions of this issue treat Black/White differences as aspects of a uniquely "American dilemma" (Myrdal, 1944). The fact is, however, that comparably disadvantaged groups exist in many countries: the Maori in New Zealand, scheduled castes ("untouchables") in India, non-European Jews in Israel, the Burakumin in Japan. All these are "caste-like" (Ogbu, 1978) or "involuntary" (Ogbu, 1994) minorities. John Ogbu distinguishes this status from that of "autonomous" minorities who are not politically or economically subordinated (like Amish or Mormons in the United States), and from that of "immigrant" or "voluntary" minorities who initially came to their new homes with positive expectations. Immigrant minorities expect their situations to improve; they tend to compare themselves favorably with peers in the old country, not unfavorably with members of the dominant majority. In contrast, to be born into a caste-like minority is to grow up firmly convinced that one's life will eventually be restricted to a small and poorly-rewarded set of social roles.

Distinctions of caste are not always linked to perceptions of race. In some countries lower and upper caste groups differ by appearance and are assumed to be racially distinct; in others they are not. The social and educational consequences are the same in both cases. All over the world, the children of caste-like minorities do less well in school than upper-caste children and drop out sooner. Where there are data, they have usually been found to have lower test scores as well.

In explaining these findings, Ogbu (1978) argues that the children of caste-like minorities do not have "effort optimism," i.e., the conviction that hard work (especially hard schoolwork) and serious commitment on their part will actually be rewarded. As a result they ignore or reject the forms of learning that are offered in school. Indeed they may practice a sort of cultural inversion, deliberately rejecting certain behaviors (such as academic achievement or other forms of "acting White") that are seen as characteristic of the dominant group. While the extent to which the attitudes described by Ogbu (1978, 1994) are responsible for African American test scores and school achievement has not been empirically established, it does seem that familiar problems can take on quite a different look when they are viewed from an international perspective.

African American culture. According to Boykin (1986, 1994), there is a fundamental conflict between certain aspects of African American culture on the one hand and the implicit cultural commitments of most American schools on the other. "When children are ordered to do their own work, arrive at their own individual answers, work only with their own materials, they are being sent cultural messages. When children come to believe that getting up and moving about the classroom is inappropriate, they are being sent powerful cultural messages. When children come to confine their 'learning' to consistently bracketed time periods, when they are consistently prompted to tell what they know and not how they feel, when they are led to believe that they are completely responsible for their own success and failure, when they are required to consistently put forth considerable effort for effort's sake on tedious and personally irrelevant tasks . . . then they are pervasively having cultural lessons imposed on them" (1994, p. 125).

In Boykin's view, the combination of constriction and competition that most American schools demand of their pupils conflicts with certain themes in the "deep structure" of African American culture. That culture includes an emphasis on such aspects of experience as spirituality, harmony, movement, verve, affect, expressive individualism, communalism, orality, and a socially defined time perspective (Boykin, 1986, 1994). While it is not shared by all African Americans to the same degree, its accessibility and familiarity give it a profound influence.

The result of this cultural conflict, in Boykin's view, is that many Black children become alienated from both the process and the products of the education to which they are exposed. One aspect of that process, now an intrinsic aspect of the culture of most American schools, is the psychometric enterprise itself. He argues (Boykin, 1994) that the successful education of African American children will require an approach that is less concerned with talent sorting and assessment, more concerned with talent development.

One further factor should not be overlooked. Only a single generation has passed since the Civil Rights movement opened new doors for African Americans, and many forms of discrimination are still all too familiar in their experience today. Hard enough to bear in its own right, discrimination is also a sharp reminder of a still more intolerable past. It would be rash indeed to assume that those experiences, and that historical legacy, have no impact on intellectual development.

The genetic hypothesis. It is sometimes suggested that the Black/White differential in psychometric intelligence is partly due to genetic differences (Jensen, 1972). There is not much direct evidence on this point, but what little there is fails to support the genetic hypothesis. One piece of evidence comes from a study of the children of American soldiers stationed in Germany after the Second World War (Eyferth, 1961): there was no mean difference between the test scores of those children whose fathers were White and those whose fathers were Black. (For a discussion of possible confounds in this study, see Flynn, 1980.) Moreover, several studies have used blood-group methods to estimate the degree of African ancestry of American Blacks; there were no significant correlations between those estimates and IQ scores (Loehlin, Vandenberg, & Osborne, 1973; Scarr, Pakstis, Katz, & Barker, 1977).

It is clear (Section 3) that genes make a substantial contribution to individual differences in intelligence test scores, at least in the White population. The fact is, however, that the high heritability of a trait within a given group has no necessary implications for the source of a difference between groups (Loehlin et al., 1975). This is now generally understood (e.g., Herrnstein & Murray, 1994). But even though no such implication is *necessary,* some have argued that a high value of h^2 makes a genetic contribution to group differences more *plausible.* Does it?

That depends on one's assessment of the actual difference between the two environments. Consider Lewontin's (1970) well-known example of seeds from the same genetically variable stock that are planted in two different fields. If the plants in field X are fertilized appropriately while key nutrients are withheld from those in field Y, we have produced an entirely environmental group difference. This example works (i.e., h^2 is genuinely irrelevant to the differential between the fields) because the differences between the effective environments of X and Y are both large and consistent. Are the environmental and cultural situations of American Blacks and Whites also substantially and consistently different—different enough to make this a good analogy? If so, the within-group heritability of IQ scores is irrelevant to the issue. Or are those situations similar enough to suggest that the analogy is inappropriate, and that one can plausibly generalize from within-group heritabilities? Thus the issue ultimately comes down to a personal judgment: How different are the relevant life experiences of Whites and Blacks in the United States today? At present, this question has no scientific answer.

6. Summary and Conclusions

Because there are many ways to be intelligent, there are also many conceptualizations of intelligence. The most influential approach, and the one that has generated the most systematic research, is based on psychometric testing. This tradition has produced a substantial body of knowledge, though many questions remain unanswered. We know much less about the forms of intelligence that tests do not easily assess: wisdom, creativity, practical knowledge, social skill, and the like.

Psychometricians have successfully measured a wide range of abilities, distinct from one another and yet intercorrelated. The complex relations among those abilities can be described in many ways. Some theorists focus on the variance that all such abilities have in common, which

Spearman termed *g* ("general intelligence"); others prefer to describe the same manifold with a set of partially independent factors; still others opt for a multifactorial description with factors hierarchically arranged and something like *g* at the top. Standardized intelligence test scores ("IQs"), which reflect a person's standing in relation to his or her age cohort, are based on tests that tap a number of different abilities. Recent studies have found that these scores are also correlated with information processing speed in certain experimental paradigms (choice reaction time, inspection time, evoked brain potentials, etc.), but the meaning of those correlations is far from clear.

Intelligence test scores predict individual differences in school achievement moderately well, correlating about .50 with grade point average and .55 with the number of years of education that individuals complete. In this context the skills measured by tests are clearly important. Nevertheless, population levels of school achievement are not determined solely or even primarily by intelligence or any other individual-difference variable. The fact that children in Japan and Taiwan learn much more mathematics than their peers in America, for example, can be attributed primarily to differences in culture and schooling rather than in abilities measured by intelligence tests.

Test scores also correlate with measures of accomplishment outside of school, e.g., with adult occupational status. To some extent those correlations result directly from the tests' link with school achievement and from their roles as "gatekeepers." In the United States today, high test scores and grades are prerequisites for entry into many careers and professions. This is not quite the whole story, however: a significant correlation between psychometric intelligence and occupational status remains even when measures of education and family background have been statistically controlled. There are also modest (negative) correlations between intelligence test scores and certain undesirable behaviors such as juvenile crime. Those correlations are necessarily low: all social outcomes result from complex causal webs in which psychometric skills are only one factor.

Like every trait, intelligence is the joint product of genetic and environmental variables. Gene action always involves a (biochemical or social) environment; environments always act via structures to which genes have contributed. Given a trait on which individuals vary, however, one can ask what fraction of that variation is associated with differences in their genotypes (this is the *heritability* of the trait) as well as what fraction is associated with differences in environmental experience. So defined, heritability (h^2) can and does vary from one population to another. In the case of IQ, h^2 is markedly lower for children (about .45) than for adults (about .75). This means that as children grow up, differences in test scores tend increasingly to reflect differences in genotype and in individual life experience rather than differences among the families in which they were raised.

The factors underlying that shift—and more generally the pathways by which genes make their undoubted contributions to individual differences in intelligence—are largely unknown. Moreover, the environmental contributions to those differences are almost equally mysterious. We know that both biological and social aspects of the environment are important for intelligence, but we are a long way from understanding how they exert their effects.

One environmental variable with clear-cut importance is the presence of formal schooling. Schools affect intelligence in many ways, not only by transmitting specific information but by developing certain intellectual skills and attitudes. Failure to attend school (or attendance at very poor schools) has a clear negative effect on intelligence test scores. Preschool programs and similar interventions often have positive effects, but in most cases the gains fade when the program is over.

A number of conditions in the biological environment have clear negative consequences for intellectual development. Some of these—very important when they occur—nevertheless do not contribute much to the population variance of IQ scores because they are relatively rare. (Perinatal complications are one such factor.) Exposure to environmental lead has well-documented negative effects; so too does prenatal exposure to high blood levels of alcohol. Malnutrition in childhood is another negative factor for intelligence, but the level at which its effects become significant has not been clearly established. Some studies suggest that dietary supplements of certain micro-nutrients can produce gains even in otherwise well-nourished individuals, but the effects are still controversial and there has been no long-term follow-up.

One of the most striking phenomena in this field is the steady worldwide rise in test scores, now often called the "Flynn effect." Mean IQs have increased more than 15 points—a full standard deviation—in the last 50 years, and the rate of gain may be increasing. These gains may result from improved nutrition, cultural changes, experience with testing, shifts in schooling or child-rearing practices, or some other factor as yet unknown.

Although there are no important sex differences in overall intelligence test scores, substantial differences do appear for specific abilities. Males typically score higher on visual-spatial and (beginning in middle childhood) mathematical skills; females excel on a number of verbal measures. Sex hormone levels are clearly related to some of these differences, but social factors presumably play a role as well. As for all the group differences reviewed here, the range of performance within each group is much larger than the mean difference between groups.

Because ethnic differences in intelligence reflect complex patterns, no overall generalization about them is appropriate. The mean IQ scores of Chinese and Japanese Americans, for example, differ little from those of Whites though their spatial ability scores tend to be somewhat higher. The outstanding record of these groups in terms of school achievement and occupational status evidently reflects cultural factors. The mean intelligence test scores of Hispanic Americans are somewhat lower than those of Whites, in part because Hispanics are often less familiar with English. Nevertheless, their test scores, like those of African Americans, are reasonably good predictors of school and college achievement.

African American IQ scores have long averaged about 15 points below those of Whites, with correspondingly lower scores on academic achievement tests. In recent years the achievement-test gap has narrowed appreciably. It is possible that the IQ-score differential is narrowing as well, but this has not been clearly established. The cause of that differential is not known; it is apparently not due to any simple form of bias in the content or administration of the tests themselves. The Flynn effect shows that environmental factors can produce differences of at least this magnitude, but that effect is mysterious in its own right. Several culturally-based explanations of the Black/White IQ differential have been proposed; some are plausible, but so far none has been conclusively supported. There is even less empirical support for a genetic interpretation. In short, no adequate explanation of the differential between the IQ means of Blacks and Whites is presently available.

It is customary to conclude surveys like this one with a summary of what has been established. Indeed, much is now known about intelligence. A near-century of research, most of it based on psychometric methods, has produced an impressive body of findings. Although we have tried to do justice to those findings in this report, it seems appropriate to conclude on a different note. In this contentious arena, our most useful role may be to remind our readers that many of the critical questions about intelligence are still unanswered. Here are a few of those questions:

1. Differences in genetic endowment contribute substantially to individual differences in (psychometric) intelligence, but the pathway by which genes produce their effects is still unknown. The impact of genetic differences appears to increase with age, but we do not know why.

2. Environmental factors also contribute substantially to the development of intelligence, but we do not clearly understand what those factors are or how they work. Attendance at school is certainly important, for example, but we do not know what aspects of schooling are critical.

3. The role of nutrition in intelligence remains obscure. Severe childhood malnutrition has clear negative effects, but the hypothesis that particular "micronutrients" may affect intelligence in otherwise adequately-fed populations has not yet been convincingly demonstrated.

4. There are significant correlations between measures of information-processing speed and psychometric intelligence, but the overall pattern of these findings yields no easy theoretical interpretation.

5. Mean scores on intelligence tests are rising steadily. They have gone up a full standard deviation in the last 50 years or so, and the rate of gain may be increasing. No one is sure why these gains are happening or what they mean.

6. The differential between the mean intelligence test scores of Blacks and Whites (about one standard deviation, although it may be diminishing) does not result from any obvious biases in test construction and administration, nor does it simply reflect differences in socioeconomic status. Explanations based on factors of caste and culture may be appropriate, but so far have little direct empirical support. There is certainly no such support for a genetic interpretation. At present, no one knows what causes this differential.

7. It is widely agreed that standardized tests do not sample all forms of intelligence. Obvious examples include creativity, wisdom, practical sense, and social sensitivity; there are surely others. Despite the importance of these abilities we know very little about them: how they develop, what factors influence that development, how they are related to more traditional measures.

In a field where so many issues are unresolved and so many questions unanswered, the confident tone that has characterized most of the debate on these topics is clearly out of place. The study of intelligence does not need politicized assertions and recriminations; it needs self-restraint, reflection, and a great deal more research. The questions that remain are socially as well as scientifically important. There is no reason to think them unanswerable, but finding the answers will require a shared and sustained effort as well as the commitment of substantial scientific resources. Just such a commitment is what we strongly recommend.

REFERENCES

Baghurst, P. A., McMichael, A. J., Wigg, N. R., Vimpani, G. V., Robertson, E. F., Roberts, R. J., & Tong, S. -L. (1992). Environmental exposure to lead and children's intelligence at the age of seven years: The Port Pirie cohort study. *New England Journal of Medicine, 327,* 1279–1284.

Bates, T. C., & Eysenck, H. J. (1993). Intelligence, inspection time, and decision time. *Intelligence, 17,* 523–531.

Baumrind, D. (1993). The average expectable environment is not good enough: A response to Scarr. *Child Development, 64,* 1299–1317.

Benbow, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. *Behavioral and Brain Sciences, 11,* 169–232.

Berenbaum, S. A., & Hines, M. (1992). Early androgens are related to childhood sex-typed toy preferences. *Psychological Science, 3,* 203–206.

Berry, J. W. (1974). Ecological and cultural factors in spatial perceptual development. In J. W. Berry & P. R. Dasen (Eds.), *Culture and cognition: Readings in cross-cultural psychology* (pp. 129–140). London: Methuen.

Bornstein, M. H. (1989). Stability in early mental development: From attention and information processing in infancy to language and cognition in childhood. In M. H. Bornstein & N. A. Krasnegor (Eds.), *Stability and continuity in mental development* (pp. 147–170). Hillsdale, NJ: Erlbaum.

Bouchard, T. J., Jr., Lykken, D. T., McGue, M., Segal, N. L., & Tellegen, A. (1990). Sources of human psychological differences: The Minnesota study of twins reared apart. *Science, 250,* 223–228.

Boykin, A. W. (1986). The triple quandary and the schooling of Afro-American children. In U. Neisser (Ed.), *The school achievement of minority children* (pp. 57–92). Hillsdale, NJ: Erlbaum.

Boykin, A. W. (1994). Harvesting talent and culture: African-American children and educational reform. In R. Rossi (Ed.), *Schools and students at risk* (pp. 116–138). New York: Teachers College Press.

Brandt, E. A. (1984). The cognitive functioning of American Indian children: A critique of McShane and Plas. *School Psychology Review, 13,* 74–82.

Brody, N. (1992). *Intelligence* (2nd ed.). San Diego, CA: Academic Press.
Broman, S. H., Nichols, P. L., & Kennedy, W. A. (1975). *Preschool IQ: Prenatal and early developmental correlates.* Hillsdale, NJ: Erlbaum.
Brown, A. L., & French, A. L. (1979). The zone of potential development: Implications for intelligence testing in the year 2000. In R. J. Sternberg & D. K. Detterman (Eds.), *Human intelligence: Perspectives on its theory and measurement* (pp. 217–235). Norwood, NJ: Ablex.
Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development, 65,* 684–698.
Carraher, T. N., Carraher, D., & Schliemann, A. D. (1985). Mathematics in the streets and in schools. *British Journal of Developmental Psychology, 3,* 21–29.
Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge, England: University of Cambridge Press.
Caryl, P. G. (1994). Early event-related potentials correlate with inspection time and intelligence. *Intelligence, 18,* 15–46.
Ceci, S. J. (1990). *On intelligence . . . more or less: A bioecological treatise on intellectual development.* Englewood Cliffs, NJ: Prentice Hall.
Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology, 27,* 703–722.
Ceci, S. J., & Liker, J. (1986). A day at the races: A study of IQ, expertise, and cognitive complexity. *Journal of Experimental Psychology: General, 115,* 255–266.
Chaiken, S. R. (1993). Two models for an inspection time paradigm: Processing distraction and processing speed versus processing speed and asymptotic strength. *Intelligence, 17,* 257–283.
Chipuer, H. M., Rovine, M., & Plomin, R. (1990). LISREL modelling: Genetic and environmental influences on IQ revisited. *Intelligence, 14,* 11–29.
Christal, R. E., Tirre, W., & Kyllonen, P. (1984). Two for the money: Speed and level scores from a computerized vocabulary test. In G. Lee & T. Ulrich (Eds.), *Proceedings, Psychology in the Department of Defense, Ninth Annual Symposium* (USAFA TR 8-2). Colorado Springs, CO: U.S. Air Force Academy.
Christiansen, K., & Knussmann, R. (1987). Sex hormones and cognitive functioning in men. *Neuropsychobiology, 18,* 27–36.
Columbo, J. (1993). *Infant cognition: Predicting later intellectual functioning.* Newbury Park, CA: Sage.
Consortium for Longitudinal Studies. (1983). *As the twig is bent . . . lasting effects of preschool programs.* Hillsdale, NJ: Erlbaum.
Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods.* New York: Irvington.
Darlington, R. B. (1986). Long-term effects of preschool programs. In U. Neisser (Ed.), *The school achievement of minority children* (pp. 159–167). Hillsdale, NJ: Erlbaum.
Davis, C., Haub, C., & Willette, J. (1983). U.S. Hispanics: Changing the face of America. *Population Bulletin, 38*(No. 3).
Deary, I. J. (1993). Inspection time and WAIS-R IQ subtypes: A confirmatory factor analysis study. *Intelligence, 17,* 223–236.
Deary, I. J. (1995). Auditory inspection time and intelligence: What is the causal direction? *Developmental Psychology, 31,* 237–250.
Driver, H. E. (1969). *Indians of North America.* Chicago: University of Chicago Press.
Dube, E. F. (1982). Literacy, cultural familiarity, and "intelligence" as determinants of story recall. In U. Neisser (Ed.), *Memory observed: Remembering in natural contexts* (pp. 274–292). New York: Freeman.
Duran, R. P. (1983). *Hispanics' education and background: Prediction of college achievement.* New York: College Entrance Examination Board.
Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly, 11,* 135–172.
Eyde, L. D. (1992). Introduction to the testing of Hispanics in industry and research. In K. F. Geisinger (Ed.), *Psychological testing of Hispanics* (pp. 167–172). Washington, DC: American Psychological Association.
Eyferth, K. (1961). Leistungen verchiedener Gruppen von Besatzungskindern im Hamburg-Wechsler Intelligentztest fur Kinder (HAWIK) [The performance of different groups of occupation children in the Hamburg-Wechsler Intelligence Test for Children]. *Archive fur die gesamte Psychologie, 113,* 222–241.
Eysenck, H. (1973). *The measurement of intelligence.* Baltimore: Williams & Wilkins.
Eysenck, H. J. (1986). Inspection time and intelligence: A historical introduction. *Personality and Individual Differences, 7,* 603–607.
Feuerstein, R. (1980). *Instrumental enrichment: An intervention program for cognitive modifiability.* Baltimore: University Park Press.
Flynn, J. R. (1980). *Race, IQ, and Jensen.* London: Routledge & Kegan Paul.
Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95,* 29–51.
Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101,* 171–191.
Flynn, J. R. (1991). *Asian-Americans: Achievement beyond IQ.* Hillsdale, NJ: Erlbaum.
Frearson, W. M., & Eysenck, H. J. (1986). Intelligence, reaction time [RT], and a new "odd-man-out" RT paradigm. *Personality and Individual Differences, 7,* 807–817.
Freedman, D. G., & Freedman, N. C. (1969). Behavioral differences between Chinese-American and European-American newborns. *Nature, 224,* 1227.
Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.
Gay, J., & Cole, M. (1967). *The new mathematics and an old culture: A study of learning among the Kpelle of Liberia.* New York: Holt, Rhinehart & Winston.
Geschwind, N., & Galaburda, A. M. (1987). *Cerebral lateralization: Biological mechanisms, associations, and pathology.* Cambridge, MA: MIT Press.
Gordon, H. W., & Lee, P. (1986). A relationship between gonadotropins and visuospatial function. *Neuropsychologia, 24,* 563–576.
Gottfried, A. W. (Ed.). (1984). *Home environment and early cognitive development: Longitudinal research.* New York: Academic Press.
Gouchie, C., & Kimura, D. (1991). The relationship between testosterone levels and cognitive ability patterns. *Psychoneuroendocrinology, 16,* 323–334.
Gould, S. J. (1978). Morton's ranking of races by cranial capacity: Unconscious manipulation of data may be a scientific norm. *Science, 200,* 503–509.
Green, R. L., Hoffman, L. T., Morse, R., Hayes, M. E., & Morgan, R. F. (1964). *The educational status of children in a district without public schools* (Cooperative Research Project No. 2321). Washington, DC: Office of Education, U.S. Department of Health, Education, and Welfare.
Grissmer, D. W., Kirby, S. N., Berends, M., & Williamson, S. (1994). *Student achievement and the changing American family.* Santa Monica, CA: RAND Corporation.
Gustafsson, J.-E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence, 8,* 179–203.
Haier, R. J. (1993). Cerebral glucose metabolism and intelligence. In P. A. Vernon (Ed.), *Bioloqical approaches to the study of human intelligence* (pp. 317–332). Norwood, NJ: Ablex.
Halpern, D. (1992). *Sex differences in cognitive abilities* (2nd ed.). Hillsdale, NJ: Erlbaum.
Halpern, D. F., & Cass, M. (1994). Laterality, sexual orientation, and immune system functioning: Is there a relationship? *International Journal of Neuroscience, 77,* 167–180.
Hart, B., & Risley, T. R. (1992). American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Developmental Psychology, 28,* 1096–1105.
Hart, B., & Risley, T. R. (in press). *Meaningful differences in the everyday experience of young American children.* Baltimore: P. H. Brookes.
Hay, D. A. (1985). *Essentials of behavior genetics.* Melbourne, Australia: Blackwell.
Heath, S. B. (1983). *Ways with words.* New York: Cambridge University Press.
Held, J. D., Alderton, D. E., Foley, P. P., & Segall, D. O. (1993). Arithmetic reasoning gender differences: Explanations found in the Armed Services Vocational Aptitude Battery (ASVAB). *Learning and Individual Differences, 5,* 171–186.

Helms, J. E. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist, 47,* 1083–1101.

Hennessy, J. J., & Merrifield, P. R. (1978). Ethnicity and sex distinctions in patterns of aptitude factor scores in a sample of urban high school seniors. *American Educational Research Journal, 15,* 385–389.

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

Herrnstein, R. J., Nickerson R. S., de Sanchez, M., & Swets, J. A. (1986). Teaching thinking skills. *American Psychologist, 41,* 1279–1289.

Hines, M. (1990). Gonadal hormones and human cognitive development. In J. Balthazart (Ed.), *Hormones, brains, and behaviors in vertebrates: 1. Sexual differentiation, neuroanatomical aspects, neurotransmitters, and neuropeptides* (pp. 51–63). Basel, Switzerland: Karger.

Hines, M., Chiu, L., McAdams, L. A., Bentler, M. P., & Lipcamon, J. (1992). Cognition and the corpus callosum: Verbal fluency, visuospatial ability, language lateralization related to midsagittal surface areas of the corpus callosum. *Behavioral Neuroscience, 106,* 3–14.

Hunt, E. (1978). Mechanics of verbal ability. *Psychological Review, 85,* 109–130.

Hunter, J. E. (1983). A causal analysis of cognitive ability, job knowledge, job performance, and supervisor ratings. In F. Landy, S. Zedeck, & J. Cleveland (Eds.), *Performance measurement and theory* (pp. 257–266). Hillsdale, NJ: Erlbaum.

Hyde, J., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107,* 139–155.

Jackson, G. D. (1975). On the report of the Ad Hoc Committee on Educational Uses of Tests with Disadvantaged Students: Another psychological view from the Association of Black Psychologists. *American Psychologist, 30,* 88–93.

Jackson, J. F. (1993). Human behavioral genetics, Scarr's theory, and her views on interventions: A critical review and commentary on their implications for African American children. *Child Development, 64,* 1318–1332.

Jackson, M., & McClelland, J. (1979). Processing determinants of reading speed. *Journal of Experimental Psychology: General, 108,* 151–181.

Janowsky, J. S., Oviatt, S. K., & Orwoll, E. S. (1994). Testosterone influences spatial cognition in older men. *Behavioral Neuroscience, 108,* 325–332.

Jardine, R., & Martin, N. G. (1983). Spatial ability and throwing accuracy. *Behavior Genetics, 13,* 331–340.

Jencks, C. (1979). *Who gets ahead? The determinants of economic success in America.* New York: Basic Books.

Jensen, A. R. (1972). *Genetics and education.* New York: Harper & Row.

Jensen, A. R. (1977). Cumulative deficit in IQ of Blacks in the rural South. *Developmental Psychology, 13,* 184–191.

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

Jensen, A. R. (1985). The nature of the black-white difference on various psychometric tests: Spearman's hypothesis. *Behavioral and Brain Sciences, 8,* 193–263.

Jensen, A. R. (1987). Individual differences in the Hick paradigm. In P. A. Vernon (Ed.), *Speed of information processing and intelligence* (pp. 101–175). Norwood, NJ: Ablex.

Jensen, A. R. (1993). Test validity: *g* vs. "tacit knowledge." *Current Directions in Psychological Science, 2,* 9–10.

Jones, H. E., and Bayley, N. (1941). The Berkeley Growth Study. *Child Development, 12,* 167–173.

Kaufman, A. S. (1994). *Intelligent testing with the WISC-III.* New York: Wiley.

Kaufman, A. S., & Doppelt, J. E. (1976). Analysis of WISC-R standardization data in terms of the stratification variables. *Child Development, 47,* 165–171.

Kohn, M. L., & Schooler, C. (1973). Occupational experience and psychological functioning: An assessment of reciprocal effects. *American Sociological Review, 38,* 97–118.

Kohn, M. L., & Schooler, C. (1983). *Work and personality: An inquiry into the impact of social stratification.* Norwood, NJ: Ablex.

Kranzler, J., & Jensen, A. R. (1989). Inspection time and intelligence: A meta-analysis. *Intelligence, 13,* 329–347.

Lave, J. (1988). *Cognition in practice.* New York: Cambridge University Press.

Law, D. J., Pellegrino, J. W., & Hunt, E. B. (1993). Comparing the tortoise and the hare: Gender differences and experience in dynamic spatial reasoning tasks. *Psychological Science, 4,* 35–40.

Leap, W. L. (1981). American Indian languages. In C. Ferguson & S. B. Heath (Eds.), *Language in the USA.* Cambridge, England: Cambridge University Press.

Lee, E. S. (1951). Negro intelligence and selective migration: A Philadelphia test of the Klineberg hypothesis. *American Sociological Review, 16,* 227–232.

Lesser, G. S., Fifer, G., & Clark, D. H. (1965). Mental abilities of children from different social-class and cultural groups. *Monographs of the Society for Research in Child Development, 30*(Whole No. 102).

Lewontin, R. (1970). Race and intelligence. *Bulletin of the Atomic Scientists, 26,* 2–8.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56,* 1479–1498.

Locurto, C. (1991). Beyond IQ in preschool programs? *Intelligence, 15,* 295–312.

Loehlin, J. C. (1989). Partitioning environmental and genetic contributions to behavioral development. *American Psychologist, 10,* 1285–1292.

Loehlin, J. C., Horn, J. M., & Willerman, L. (1989). Modeling IQ change: Evidence from the Texas Adoption Project. *Child Development, 60,* 993–1004.

Loehlin, J. C., Lindzey, G., & Spuhler, J. N. (1975). *Race differences in intelligence.* New York: Freeman.

Loehlin, J. C., Vandenberg, S. G., & Osborne, R. T. (1973). Blood group genes and Negro-White ability differences. *Behavior Genetics, 3,* 263–270.

Longstreth, L. E. (1984). Jensen's reaction-time investigations of intelligence: A critique. *Intelligence, 8,* 139–160.

Lubchenko, L. O. (1976). *The high-risk infant.* Philadelphia: Saunders.

Lynn, R. (1982). IQ in Japan and the United States shows a growing disparity. *Nature, 297,* 222–223.

Lynn, R. (1990). The role of nutrition in secular increases in intelligence. *Personality and Individual Differences, 11,* 273–285.

Lynn, R. (1993). Oriental Americans: Their IQ, educational attainment, and socio-economic status. *Personality and Individual Differences, 15,* 237–242.

Lynn, R. (1994). Sex differences in intelligence and brain size: A paradox resolved. *Personality and Individual Differences, 17,* 257–271.

Mackenzie, B., Molloy, E., Martin, F., Lovegrove, W., & McNicol, D. (1991). Inspection time and the content of simple tasks: A framework for research on speed of information processing. *Australian Journal of Psychology, 43,* 37–43.

Masters, M. S., & Sanders, B. (1993). Is the gender difference in mental rotation disappearing? *Behavior Genetics, 23,* 337–341.

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

McCartney, K., Harris, M. J., & Bernieri, F. (1990). Growing up and growing apart: A developmental meta-analysis of twin studies. *Psychological Bulletin, 107,* 226–237.

McGue, M., Bouchard, T. J., Jr., Iacono, W. G., & Lykken, D. T. (1993). Behavioral genetics of cognitive ability: A life-span perspective. In R. Plomin & G. E. McClearn (Eds.), *Nature, nurture, & psychology* (pp. 59–76). Washington, DC: American Psychological Association.

McMichael, A. J., Baghurst, P. A., Wigg, N. R., Vimpani, G. V., Robertson, E. F., & Roberts, R. J. (1988). Port Pirie cohort study: Environmental exposure to lead and children's abilities at the age of four years. *New England Journal of Medicine, 319,* 468–475.

McShane, D. A., & Berry, J. W. (1988). Native North Americans: Indian and Inuit abilities. In S. H. Irvine & J. W. Berry (Eds.), *Human abilities in cultural context* (pp. 385–426). New York: Cambridge University Press.

McShane, D. A., & Cook, V. J. (1985). Transcultural intellectual assessment: Performance by Hispanics on the Wechsler Scales. In B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurements, and applications.* New York: Wiley.

McShane, D. A., & Plas, J. M. (1984a). Response to a critique of the McShane & Plas review of American Indian performance on the Wechsler Intelligence Scales. *School Psychology Review, 13,* 83–88.
McShane, D. A., & Plas, J. M. (1984b). The cognitive functioning of American Indian children: Moving from the WISC to the WISC-R. *School Psychology Review, 13,* 61–73.
Meehan, A. M. (1984). A meta-analysis of sex differences in formal operational thought. *Child Development, 55,* 1110–1124.
Mercer, J. R. (1984). What is a racially and culturally nondiscriminatory test? A sociological and pluralistic perspective. In C. R. Reynolds & R. T. Brown (Eds.), *Perspectives on bias in mental testing.* New York: Plenum Press.
Moffitt, T. E., Caspi, A., Harkness, A. R., & Silva, P. A. (1993). The natural history of change in intellectual performance: Who changes? How much? Is it meaningful? *Journal of Child Psychology and Psychiatry, 34,* 455–506.
Moffitt, T. E., Gabrielli, W. F., Mednick, S. A., & Schulsinger, F. (1981). Socioeconomic status, IQ, and delinquency. *Journal of Abnormal Psychology, 90,* 152–156.
Myrdal, G. (1944). *An American dilemma: The Negro problem and modern democracy.* New York: Harper.
Needleman, H. L., Geiger, S. K., & Frank, R. (1985). Lead and IQ scores: A reanalysis. *Science, 227,* 701–704.
Neisser, U. (1976). General, academic, and artificial intelligence. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 135–144). Hillsdale, NJ: Erlbaum.
Nettelbeck, T. (1987). Inspection time and intelligence. In P. A. Vernon (Ed.), *Speed of information-processing and intelligence* (pp. 295–346). Norwood, NJ: Ablex.
Nyborg, H. (1984). Performance and intelligence in hormonally different groups. In G. J. DeVries, J. DeBruin, H. Uylings, & M. Cormer (Eds.), *Progress in brain research* (Vol. 61, pp. 491–508). Amsterdam: Elsevier Science.
Ogbu, J. U. (1978). *Minority education and caste: The American system in cross-cultural perspective.* New York: Academic Press.
Ogbu, J. U. (1994). From cultural differences to differences in cultural frames of reference. In P. M. Greenfield & R. R. Cocking (Eds.), *Cross-cultural roots of minority child development* (pp. 365–391). Hillsdale, NJ: Erlbaum.
Okagaki, L., & Sternberg, R. J. (1993). Parental beliefs and children's school performance. *Child Development, 64,* 36–56.
Olivier, G. (1980). The increase of stature in France. *Journal of Human Evolution, 9,* 645–649.
Pascual-Leone, J., & Ijaz, H. (1989). Mental capacity testing as form of intellectual-developmental assessment. In R. J. Samuda, S. L. Kong, et al. (Eds.), *Assessment and placement of minority students.* Toronto, Ontario, Canada: Hogrefe & Huber.
Pedersen, N. L., Plomin, R., Nesselroade, J. R., & McClearn, G. E. (1992). A quantitative genetic analysis of cognitive abilities during the second half of the life span. *Psychological Science, 3,* 346–353.
Pennock-Roman, M. (1992). Interpreting test performance in selective admissions for Hispanic students. In K. F. Geisinger (Ed.), *Psychological testing of Hispanics* (pp. 95–135). Washington, DC: American Psychological Association.
Piaget, J. (1972). *The psychology of intelligence.* Totowa, NJ: Littlefield Adams.
Plomin, R., & Bergeman, C. S. (1991). The nature of nurture: Genetic influence on "environmental" measures. *Behavioral and Brain Sciences, 14,* 373–427.
Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin, 84,* 309–322.
Plomin, R., DeFries, J. C., & McClearn (1990). *Behavioral genetics: A primer* (2nd ed.). New York: Freeman.
Pollitt, E., Gorman, K. S., Engle, P. L., Martorell, R., & Rivera, J. (1993). Early supplementary feeding and cognition. *Monographs of the Society for Research in Child Development, 58*(Serial No. 235).
Posner, M. I., Boies, S. J., Eichelman, W. H., & Taylor, R. L. (1969). Retention of visual and name codes of single letters. *Journal of Experimental Psychology, 79,* 1–16.
Reed, T. E., & Jensen, A. R. (1992). Conduction velocity in a brain nerve pathway of normal adults correlates with intelligence level. *Intelligence, 16,* 259–272.
Reed, T. E., & Jensen, A. R. (1993). Choice reaction time and visual pathway conduction velocity both correlate with intelligence but appear not to correlate with each other: Implications for information processing. *Intelligence, 17,* 191–203.
Rehberg, R. A., & Rosenthal, E. R. (1978). *Class and merit in the American high school.* New York: Longman.
Resnick, S. M., Berenbaum, S. A., Gottesman, I. I., & Bouchard, T. J., Jr. (1986). Early hormonal influences on cognitive functioning in congenital adrenal hyperplasia. *Developmental Psychology, 22,* 191–198.
Reynolds, C. R., & Brown, R. T. (1984). Bias in mental testing: An introduction to the issues. In C. R. Reynolds & R. T. Brown (Eds.), *Perspectives on bias in mental testing* (pp. 1–39). New York: Plenum Press.
Reynolds, C. R., Chastain, R. L., Kaufman, A. S., & McLean, J. E. (1987). Demographic characteristics and IQ among adults: Analysis of the WAIS-R standardization sample as a function of the stratification variables. *Journal of School Psychology, 25,* 323–342.
Ricciuti, H. N. (1993). Nutrition and mental development. *Current Directions in Psychological Science, 2,* 43–46.
Roberts, R. C. (1967). Some concepts and methods in quantitative genetics. In J. Hirsch (Ed.), *Behavior-genetic analysis* (pp. 214–257). New York: McGraw-Hill.
Rodriguez, O. (1992). Introduction to technical and societal issues in the psychological testing of Hispanics. In K. F. Geisinger (Ed.), *Psychological testing of Hispanics* (pp. 11–15). Washington, DC: American Psychological Association.
Rose, R. J., Harris, E. L., Christian, J. C., & Nance, W. E. (1979). Genetic variance in non-verbal intelligence: Data from the kinships of identical twins. *Science, 205,* 1153–1155.
Rose, S. A., & Feldman, J. (1995). The prediction of IQ and specific cognitive abilities at 11 years from infancy measures. *Developmental Psychology, 31,* 685–696.
Rosetti, L. (1986). *High risk infants: Identification, assessment, and intervention.* Boston: Little Brown.
Rush, D., Stein, Z., Susser, M., & Brody, N. (1980). Outcome at one year of age: Effects on somatic and psychological measures. In D. Rush, Z. Stein, & M. Susser (Eds.), *Diet in pregnancy: A randomized controlled trial of nutritional supplements.* New York: Liss.
Saudino, K. J., Plomin, R., Pedersen, N. L., & McClearn, G. E. (1994). The etiology of high and low cognitive ability during the second half of the life span. *Intelligence, 19,* 359–371.
Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Development, 63,* 1–19.
Scarr, S. (1993). Biological and cultural diversity: The legacy of Darwin for development. *Child Development, 64,* 1333–1353.
Scarr, S., Pakstis, A. J., Katz, S. H., & Barker, W. B. (1977). Absence of a relationship between degree of White ancestry and intellectual skills within a Black population. *Human Genetics, 39,* 69–86.
Scarr, S., & Weinberg, R. A. (1978). The influence of "family background" on intellectual attainment. *American Sociological Review, 43,* 674–692.
Scarr, S., Weinberg, R. A., & Waldman, I. D. (1993). IQ correlations in transracial adoptive families. *Intelligence, 17,* 541–555.
Schmidt, F. L., & Hunter, J. E. (1993). Tacit knowledge, practical intelligence, and job knowledge. *Current Directions in Psychological Science, 2,* 8–9.
Schneider, B., Hieshima, J. A., Lee, S., & Plank, S. (1994). East-Asian academic success in the United States: Family, school, and cultural explanations. In P. M. Greenfield & R. R. Cocking (Eds.), *Cross-cultural roots of minority child development* (pp. 332–350). Hillsdale, NJ: Erlbaum.
Schoenthaler, S. J., Amos, S. P., Eysenck, H. J., Peritz, E., & Yudkin, J. (1991). Controlled trial of vitamin-mineral supplementation: Effects on intelligence and performance. *Personality and Individual Differences, 12,* 351–362.
Seashore, H., Wesman, A., & Doppelt, J. (1950). The standardization of the Wechsler Intelligence Scale for Children. *Journal of Consulting Psychology, 14,* 99–110.

Serpell, R. (1974). *Estimates of intelligence in a rural community of Eastern Zambia: Human Development Research Unit Reports, 25.* Mimeo, Lusaka: University of Zambia.

Serpell, R. (1979). How specific are perceptual skills? A cross-cultural study of pattern reproduction. *British Journal of Psychology, 70,* 365–380.

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Constable, R. T., Skudlarski, P., Fulbright, R. K., Bronen, R. A., Fletcher, J. M., Shankweller, D. P., Katz, L., & Gore, J. C. (1995). Sex differences in the functional organization of the brain for language. *Nature, 373,* 607–609.

Sigman, M. (1995). Nutrition and child development: More food for thought. *Current Directions in Psychological Science, 4,* 52–55.

Spearman, C. (1927). *The abilities of man.* New York: Macmillan.

Stanley, J. (1993). Boys and girls who reason well mathematically. In G. R. Bock and K. Ackrill (Eds.), *The origins and development of high ability.* Chichester, England: Wiley.

Stanley, J. C., Benbow, C. P., Brody, L. E., Dauber, S., & Lupkowski, A. (1992). Gender differences on eighty-six nationally standardized aptitude and achievement tests. In N. Colangelo, S. G. Assouline, & D. L. Ambroson (Eds.), *Talent development, Vol. 1: Proceedings from the 1991 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development.* Unionville, NY: Trillium Press.

Stein, Z., Susser, M., Saenger, G., & Marolla, F. (1975). *Famine and human development: The Dutch hunger winter of 1944–45.* New York: Oxford University Press.

Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities.* Hillsdale, NJ: Erlbaum.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence.* New York: Cambridge University Press.

Sternberg, R. J. (Ed.). (1994). *Encyclopedia of human intelligence.* New York: MacMillan.

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

Sternberg, R. J., & Wagner, R. K. (1993). The geocentric view of intelligence and job performance is wrong. *Current Directions in Psychological Science, 2,* 1–4.

Sternberg, R. J., Wagner, R. K., Williams, W. M., & Horvath, J. A. (1995). Testing common sense. *American Psychologist, 50,* 912–927.

Stevenson, H. W., Lee, S. Y., & Stigler, J. W. (1986). Mathematics achievement of Chinese, Japanese, and American children. *Science, 231,* 693–699.

Stevenson, H. W., & Stigler, J. W. (1992). *The learning gap.* New York: Summit Books.

Stevenson, H. W., Stigler, J. W., Lee, S. Y., Lucker, G. W., Kitamura, S., & Hsu, C. C. (1985). Cognitive performance and academic achievement of Japanese, Chinese, and American children. *Child Development, 56,* 718–734.

Streissguth, A. P., Barr, H. M., Sampson, P. D., Darby, B. L., & Martin, D. C. (1989). IQ at age 4 in relation to maternal alcohol use and smoking during pregnancy. *Developmental Psychology, 25,* 3–11.

Super, C. M. (1983). Cultural variation in the meaning and uses of children's "intelligence." In J. B. Deregowski, S. Dziurawiec, & R. C. Annis (Eds.), *Explorations in cross-cultural psychology.* Lisse, The Netherlands: Swets & Zeitlinger.

Sutaria, S. D. (1985). *Specific learning disabilities: Nature and needs.* Springfield, IL: Charles C Thomas.

Terman, L. M. (1925). *Genetic studies of genius: Mental and physical traits of a thousand gifted children.* Stanford, CA: Stanford University Press.

Terman, L. M., & Merrill, M. A. (1937). *Measuring intelligence: A guide to the administration of the new revised Stanford-Binet tests of intelligence.* Boston: Houghton Mifflin.

Thomson, G. H. (1939). *The factorial analysis of human ability.* Boston: Houghton Mifflin.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet intelligence scale: Fourth edition (Technical Manual).* Chicago: Riverside.

Thurstone, L. L. (1938). *Primary mental abilities.* Chicago: University of Chicago Press.

Turkheimer, E. (1991). Individual and group differences in adoption studies of IQ. *Psychological Bulletin, 110,* 392–405.

United States Bureau of the Census. (1994). *The Hispanic population of the United States: March 1993* (Current Population Reports, Series P20-475). Washington, DC: Author.

United States National Commission for Employment Policy. (1982). *Hispanics and jobs: Barriers to progress* (Report No. 14). Washington, DC: Author.

Vernon, P. A. (1987). *Speed of information processing and intelligence.* Norwood, NJ: Ablex.

Vernon, P. A. (1993). *Biological approaches to the study of human intelligence.* Norwood, NJ: Ablex.

Vickers, D., Nettelbeck, T., & Wilson, R. J. (1972) Perceptual indices of performance: The measurement of "inspection time" and "noise" in the visual system. *Perception, 1,* 263–295.

Vincent, K. R. (1991). Black/White IQ differences: Does age make the difference? *Journal of Clinical Psychology, 47,* 266–270.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Waller, J. H. (1971). Achievement and social mobility: Relationships among IQ score, education, and occupation in two generations. *Social Biology, 18,* 252–259.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91,* 461–481.

Wober, M. (1974). Towards an understanding of the Kiganda concept of intelligence. In J. W. Berry & P. R. Dasen (Eds.), *Culture and cognition: Readings in cross-cultural psychology* (pp. 261–280). London: Methuen.

Yairi, E., & Ambrose, N. (1992). Onset of stuttering in preschool children: Selected factors. *Journal of Speech and Hearing Research, 35,* 782–788.