Intelligence, Part II: Validating Intelligence – Correlates of IQ (Causes and Consequences)

6

Key Terms

event-related potential
inspection time
job analysis
reaction time
Wechsler Adult Intelligence Scale (WAIS)

Chapter Outline

6.1 INTRODUCTION

Chapter 5 introduced the psychological concept of intelligence, starting with an examination of the historical development of early theories and measurement approaches, leading to salient structural issues such as the general intelligence factor (*g*), the distinction between fluid (*gf*) and crystallized (*gc*) intelligences, and the hierarchical structure of human abilities. The aim of this chapter is to follow up some of these themes by assessing the research evidence for the *validity* of ability tests, notably *g* and IQ. As with personality traits in chapter 3, then, this chapter looks at whether ability tests are useful predictors of a wide range of behavioral outcomes, including occupational and academic performance, health, and longevity. Accordingly, it addresses the question of what it means to score high on ability tests, that is, what consequences this may have and whether, when, and where it matters to be more intelligent.

Whereas some correlates of IQ are usually interpreted in terms of outcomes, other correlates are often regarded as indicators of the causes of individual differences in cognitive ability. Amongst the latter are basic information processing tasks and measures of brain efficiency, such as reaction time and electroencephalogram activity. Moreover, the study of biological differences in intelligence (also discussed in chapter 7) has included what is arguably the most controversial research area in differential psychology, namely, group differences in cognitive ability – particularly sex and race. This chapter also deals with these issues.

6.2 WECHSLER'S IQ SCALE

Before examining the salient correlates of intelligence, it is important to look at Wechsler's IQ scales, which have represented the most widely used measure of intelligence for decades. Introduced in 1939 as the Wechsler-Bellevue test, the scales progressively replaced the Stanford adaptation of Binet's test. One major reason for this was that, unlike Terman's scale, the Wechsler test could also be used to measure adult IQ (after the age of 14), and was validated on large and representative samples. For instance, the 1955 revision of this scale, relabeled the **Wechsler Adult Intelligence Scale (WAIS)**, was based on over 2,000 individuals, aged 16–75. Moreover, Wechsler designed a specific version of his test for children (aged 5–16), called the *Wechsler Intelligence Scale for Children* (WISC).

Another advantage of Wechsler's scales was that scores could be calculated and interpreted on the basis of between-subjects comparisons rather than the (mental age/chronological age) × 100 formula. Since mental age remains pretty much the same after the (chronological) age of 16, Wechsler's readjustment and standardization allowed him – and any test administrator – to compare testees' scores with an "expected" or "typical" score obtained by other testees. Sex, social class, nationality, and other group factors were carefully stratified so as to maximize accuracy in the interpretation of scores. Applied to the concept of normal distribution (see Figure 5.3 and section 5.3.3), Wechsler's formula of (actual test score/expected score) × 100 could then be used to assign test-takers a "relative" score with regard to the overall population or specific samples, representing differences in terms of standard deviation.

Wechsler Adult Intelligence Scale (WAIS) widely used measure of intelligence which has progressively replaced the Stanford/Binet test because of its suitability for measuring adult IQ; scores are calculated on the basis of between-subjects comparisons rather than on the (mental age/chronological age) × 100 formula

Although IQ is a quantification of standardized differences between individuals' performance, neither the formula nor the normal distribution or "bell curve" of scores refers to parametric data (see Figure 5.3). In contrast, IQ scores and scales are *non-parametric* in nature, which means there is no absolute zero and the distance or interval between two data points is not homogeneous. Thus, an individual with an IQ score of 100 is not twice as clever as someone with an IQ score of 50, or half as bright as someone with an IQ score of 200 (if such a person existed). Further, the difference between an IQ of 120 and one of 130 is not the same as that between an IQ of 90 and one of 100, because scores are interpreted in terms of the relative position to others. If, then, most people tend to score between 85 and 115, scores outside this range are less frequent and every point difference outside this range represents more significant differences between individuals.

The construction of Wechsler's scale was influenced not only by Terman's (1916) American version of Binet's IQ scale but also by the army-oriented scale developed by Robert Yerkes in 1919, namely, the National Intelligence Test of the United States. This test comprised two different subscales, the Alpha and Beta scales, measuring verbal and non-verbal ability, respectively. Likewise, the WAIS comprises different subscales of verbal and non-verbal (performance) scales (see Table 6.1). Verbal scales include *information*, *vocabulary*, *comprehension*, *arithmetic*, *similarities*, and *digit span*. Performance subtests comprise *picture completion*, *picture arrangement*, *block design*, *object assembly*, and *digit symbol*. The distinction between verbal and performance tests is based on empirical rather than conceptual grounds, specifically the use of factor analyses and other statistical techniques (i.e., some sections are correlated with each other, whilst others are not).

The inclusion of a wide range of subtests enabled Wechsler to measure intelligence in a global, comprehensive way, without however disregarding specific abilities. As will be noted (see chapter 8), there has been extensive debate on whether intelligence should be conceptualized as a general, single mental capacity or as a large number of unrelated abilities. The WAIS seems to represent a third-way solution, a compromise between *splitters* (those who believe there are many distinct, independent abilities) and *lumpers* (those who believe that intelligence is a general, single psychological attribute), just as Carroll's (1993) hierarchical model prescribes. In the words of Wechsler (1958): "While intelligence may manifest itself in a variety of ways, one must assume that there is some commonality or basic similarity between those forms of behavior which one identifies as intelligent" (p. 5).

Thus researchers have largely focused on general cognitive ability or IQ when validating intelligence. Although this approach can be justified on both conceptual and psychometric grounds, it

Table 6.1 WAIS structure

WAIS subtests (verbal and performance)[a]	
Verbal	*Performance*
Information: Tests knowledge on various subjects (e.g., science, history, arts)	*Picture completion*: Presents illustrations of incomplete objects and requires testees to complete them
Vocabulary: Requires testees to provide definitions for words	*Picture arrangement*: Requires the person to put a disarranged sequence of pictures/cards in order, to recreate a story
Comprehension: Tests the individual's ability to understand sayings, rules, or proverbs	*Block design*: Tests the ability to form quick patterns with cubes of different colors
Arithmetic: Mental calculations (if 15 oranges cost $3, how much will 7 oranges cost?)	*Object assembly*: Similar to block design, involves disarranged objects which make up a jigsaw
Similarities: Asks people to relate two different concepts or objects (by identifying the underlying characteristic in common)	*Digit symbol*: Requires the person to memorize specific codes for different numbers and fill in a sequence with those symbols
Digit span: Requires the person to repeat a sequence of digits read out by the examiner (both in normal and reverse order)	

[a] Correlations between different subtests range from $r = .33$ (object assembly and digit span) to $r = .81$ (vocabulary and information).

fails to provide a detailed account of the processes underlying the correlations between different test parts and why certain types of tasks are more intercorrelated than others. Accordingly, the choice of particular tests and, consequently, identification of specific aspects of intellectual ability are matters of empirical evidence: if, in a large and representative sample, there is a general tendency for people who do well in some sections of the tests to do well in others, all sections can be justifiably included as part of the scale and considered partial measures of intellectual ability.

Conversely, if a section of the test does not distinguish between individuals' performance on other sections, it should neither be included in the scale nor be considered a measure of intelligence. In that sense, it could be argued (as critics have) that the only reason why IQ tests seem to measure a single and general underlying intelligence is because the people who designed these tests have chosen to do so. Yet, the meaning and usefulness of any IQ test, as well as the very concept of intelligence, can only be judged against external indicators of validity, hence the importance of this chapter. The forthcoming sections deal with the validity of IQ and *g* as predictors of different performance and behavioral outcomes. As will be seen, intelligence is a highly pragmatic, functional variable with pervasive effects across a wide range of settings and outcomes, and individual differences in cognitive ability have clear implications in everyday life.

6.3 INTELLIGENCE AT SCHOOL AND UNIVERSITY: EDUCATIONAL OUTCOMES

To say that IQ tests predict school performance is almost tautological because ability tests were specifically designed to predict individual differences in school and educational success (see chapter 5). It is therefore unlikely that any ability test uncorrelated with school success (or learning outcomes in general) would meet the criteria for intelligence tests or be labeled "intelligence." Nonetheless, educational psychologists (and indeed some intelligence researchers) have often raised doubts about the predictive power of ability tests in academic settings. Furthermore, academic assessment methods, particularly in higher education, are increasingly focused on continuous assessment or coursework assignments, which make academic performance more dependable on personality than cognitive ability (Chamorro-Premuzic & Furnham, 2005, 2006).

That cognitive ability tests such as *g* or IQ are accurate predictors of student performance, particularly during primary and secondary school, has been replicated for over a century (e.g., Binet, 1903; Binet & Simon, 1905; Brody, 2000; Harris, 1940; Terman, 1916; Thurstone, 1919; Willingham, 1974). In fact, psychometric intelligence is by far the most robust and consistent predictor of academic performance (Elshout & Veenman, 1992; Gagne & St. Pere, 2001; Sternberg & Kaufman, 1998) and educational level in general (Brand, 1994). Some examples are summarized in Table 6.2.

However, the predictive power of cognitive ability in educational settings seems to decrease as students progress to higher academic levels, probably because of restrictions of range in intelligence (i.e., brighter students are more likely to pursue further education, making ability levels more and more homogeneous). In fact, studies have often found weak or non-significant relationships between ability and academic performance measures beyond secondary school (see Mehta & Kumar, 1985; Sanders, Osborne, & Greene, 1955; Seth & Pratap, 1971; Singh & Varma, 1995; Thompson, 1934). Even such stalwart supporters of intelligence as Jensen (1980) reported a drop in correlations from $r = .70$ in elementary school to $r = .50$ in secondary school and $r = .40$ in college (see also Boekaerts, 1995). Likewise, Hunter (1986) argued that measures of *g*, as well as verbal and quantitative abilities, have only been found to be modest predictors of academic success for adults (see Figure 6.1).

The fact that ability tests may show weakened predictive validity at higher levels of education is also consistent with the increasing significance of personality or non-cognitive traits at

Table 6.2 Intelligence and academic performance

Author (year of publication)	*Key findings*
Bright (1930)	High correlations between ability test and both academic and citizenship grades in public schools
Springsteen (1940)	Cognitive ability correlates with school performance in a sample of mentally handicapped school pupils
Tenopyr (1967)	Cognitive ability (SCAT) is a more powerful predictor of academic achievement than social intelligence (findings partly replicated by Riggio, Messamer, & Throckmorton, 1991)
Sharma & Rao (1983)	Hindu female school students' academic performance correlates with non-verbal intellectual ability (Raven's Progressive Matrices)
Bachman et al. (1986)	IQ test better predictor of primary school grades than measures of abnormal behavior
Walberg, Strykowski, Rovai, & Hung (1984)	Meta-analysis of more than 3,000 studies reported correlation in the order of $r = .70$ between the two constructs (replicated in Gagne & St. Pere, 2001)
Willingham (1974)	The graduate record examination (GRE) correlates substantially with cognitive ability and future performance at university
Kuncel, Hezlett, & Ones (2001)	A very large-scale meta-analysis ($N = 82,659$) shows strong predictive power for GRE and undergraduate grade point average (UGPA) as predictors of postgraduate education level. Yet these measures also confound non-cognitive variables such as personality traits

Source: Based on Chamorro-Premuzic & Furnham (2005).

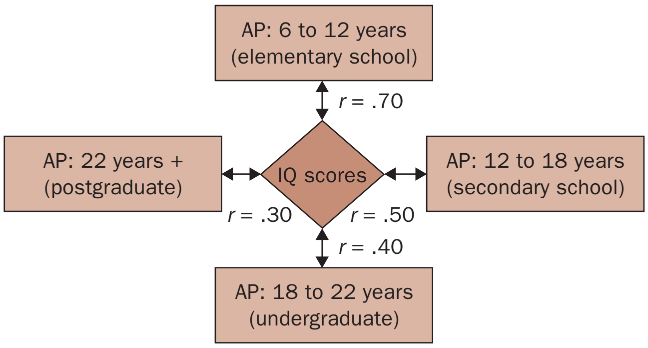

Figure 6.1 Correlations between intelligence and academic performance (AP) at different levels of education. All *r* values are approximate. *Source*: Based on Ackerman (1994); Boekaerts (1995); Hunter (1986); Jensen (1980).

such academic stages (Chamorro-Premuzic & Furnham, 2005). Indeed, researchers have recently increased the search for additional predictors of educational outcomes in the hope of explaining further variance in student attainment levels (see Ackerman & Beier, 2003; Ackerman & Heggestad, 1997). Thus Ackerman and Rolfhus (1996) argued that "abilities are only one part of the complex causal framework that determines whether a student pursues the acquisition of knowledge and skills within a particular domain. Two other components of the equation are interests and personality traits" (p. 176). See also chapters 3, 9, and 12.

6.4 IN THE JOB: OCCUPATIONAL OUTCOMES OF INTELLIGENCE

Ability tests have been used as predictors of job or occupational performance for almost a century. In fact, some of the best-known tests were developed in the context of job performance, particularly in military settings. Thus Robert Yerkes developed a purpose measure for the army (the National Intelligence Test, famous for its Alpha and Beta scales) as early as 1919, and ability tests were used to recruit and train fighter pilots in World War II (Matarazzo, 1972). Ever since, army data have represented an important source of information to assess the validity of ability tests, no doubt due to the large and representative samples they comprise. In more recent times, McHenry, Hough, Toquam, Hanson, and Ashworth (1990) published a large and comprehensive meta-analysis on the correlations between *g* (measured by the Armed Services Vocational Aptitude Battery test) and military performance. As seen in Table 6.3, *g* is substantially correlated with technical proficiency and general soldiering performance, and moderately correlated with effort and leadership as well as physical fitness and military bearing. Measures of personal discipline, on the other hand, are only modestly correlated with *g* (and arguably more dependent on personality or non-cognitive traits).

The most compelling evidence for the importance of *g* in military settings derives from a study by O'Toole and Stankov (1992), where the authors looked at the relationship between IQ scores (used in military selection) and non-combat deaths at the age of 40 in a sample that included over 2,000 Australian veterans. Even after controlling for over 50 behavioral, psychological, and health

Table 6.3 Correlations of *g* and military performance

Technical proficiency	.63
General soldiering	.65
Effort and leadership	.31
Personal discipline	.16
Physical fitness and military bearing	.20

Source: Adapted from McHenry et al. (1990).

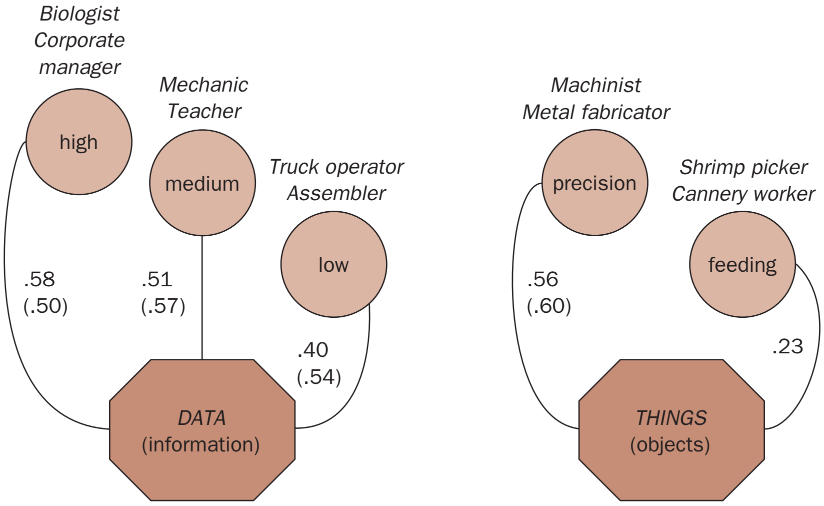

Figure 6.2 Predictive validity of cognitive ability across different job types.
All numbers are correlation coefficients; numbers in brackets refer to training performance; numbers outside brackets refer to job performance.
Source: Based on Hunter (1983); Hunter & Hunter (1984).

variables, IQ scores predicted risks of death. In fact, with every additional IQ point, there was a 15 percent decrease in death risk (see also section 6.5 on health and longevity).

It is a well-replicated finding that the more complex the job, the more important and stronger the effects of *g*. Thus the correlations between cognitive ability measures and job performance are moderated by job complexity. As one would expect, intellectually demanding jobs are substantially correlated with ability tests or *g-loaded*, whereas jobs that do not involve reasoning or intellectual tasks correlate lower with IQ.

In one of the first comprehensive meta-analyses of the relationship between intelligence and job performance, Hunter (1983; Hunter & Hunter, 1984) showed that cognitive ability, as measured psychometrically through the US Employment Service General Aptitude Test Battery, was significantly correlated with a wide variety of jobs including 515 occupations. Indeed, Hunter classified different jobs according to established norms and job complexity – a method called **job analysis** – and reported different correlations for each job family.

job analysis a method of classifying different jobs according to the nature and complexity of the work as well as the relationships of the jobholder with other people

As shown in Figure 6.2, jobs were divided into *data* and *things*, according to whether individuals were more involved in manipulating information or physical objects, respectively. In turn, both groups were further divided according to complexity, namely, *high*, *medium*, and *low* in the case of data/information, and *precision* and *feeding* in the case of things/objects. Ability measures were correlated not only with job performance but also with training performance (which refers to an individual's ability to learn the required skills and tasks quickly and accurately). This distinction is important for at least two reasons. First, individuals may not always "replicate" their training performance on the actual job, because their motivation and incentives may decrease after they start working (see chapter 9). Second, training may *mediate* the relationship between cognitive ability and job performance. This means ability levels may not only have direct effects on job performance, but also influence how quickly and well individuals will learn and be trained, which, in turn, will further affect job performance levels (see Figure 6.3).

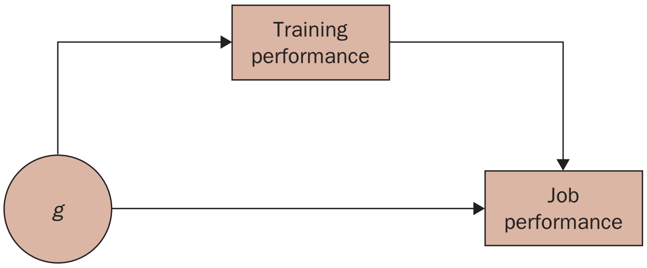

Figure 6.3 Mediated and direct effects of *g* on job performance.

As seen in Figure 6.2, correlations between ability and both job performance and training increase with job complexity. For data jobs, the correlation between ability tests and job performance is .58 when the job complexity is high, .52 when it is medium, and .40 when it is low. This pattern of results is not manifested across measures of training performance. Yet, when we look at "things" (jobs involving manipulation of objects rather than information), there is a substantial difference in the size of correlations for high (precision) and low (feeding) job complexity. Thus *g* is most important when the job is intellectually demanding and least important when the job is not intellectually demanding (as occurs in most physically demanding jobs related to manipulation of things).

Experts have also emphasized that intellectually demanding jobs are not necessarily dependent on academic expertise or high

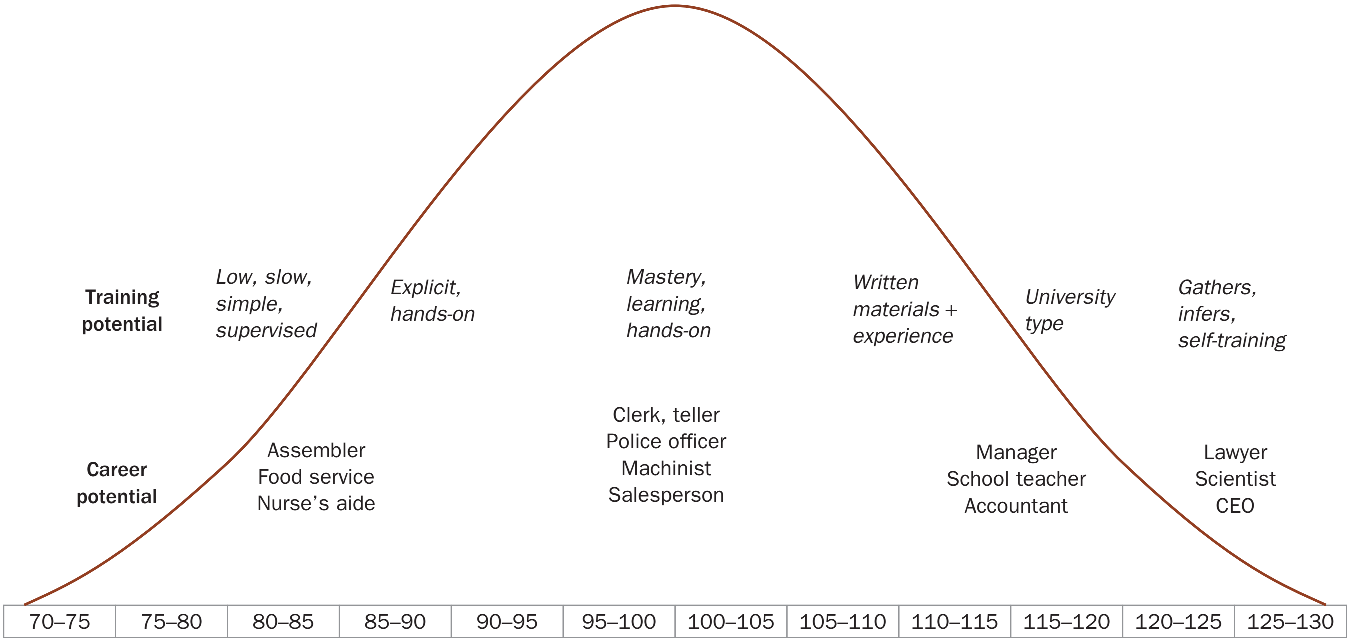

Figure 6.4 Occupational consequences of IQ.
Source: Based on Gottfredson (2004b).

educational attainment (Gottfredson, 1997). For example, the chief executive of a company may have few academic credentials and yet have one of the most intellectually demanding jobs, whilst dyslexic individuals may excel in the workforce despite failing in educational settings (see also chapter 8).

Whilst it is important to understand that IQ is not destiny (even its highest correlation with job performance leaves unexplained variance), it would be fallacious to deny its importance in the workplace. In fact, cognitive ability provides a more accurate estimate of a person's potential for the job than other psychological or non-psychological variables. More importantly, it is more objective and less exposed to bias than other methods. Figure 6.4 illustrates the typical distribution of IQ scores across a variety of occupations in terms of both training and career potential.

6.5 INTELLIGENCE, LONGEVITY, AND HEALTH

Intelligence researchers have also examined the validity of ability measures in regards to health outcomes and longevity. If intelligence represents an important adaptational tool, it should be significantly correlated with positive health outcomes as well as longevity, meaning brighter people should be generally healthier and live longer than their less bright counterparts.

Traditionally, health and even differential psychologists have emphasized the importance of motivational and non-cognitive factors such as personality traits (see section 3.6) on health outcomes. Yet, recent studies have indicated that abilities may be even more influential when it comes to predicting health and longevity. In fact, longitudinal data on the validity of IQ as a predictor of a variety of social outcomes have provided compelling evidence for the importance of *g* in real life. Gottfredson (2004a) reported associations between *g* and the following health-related outcomes:

- Physical fitness
- Low-sugar diet
- Low-fat diet
- Longevity
- Alcoholism (negative)
- Infant mortality (negative)
- Smoking (negative)
- Obesity (negative)

Whilst these behaviors are also associated with socioeconomic factors (such that poor or deprived groups tend to be more at risk than wealthy or educated individuals), Gottfredson emphasized the importance of cognitive ability over and above socioeconomic variables. In fact, it seems that the increase in availability of resources and improvements in socioeconomic conditions do little to reduce group differences between educated and less educated individuals. Rather, the more resources and information are available, the bigger the gap between lower- and higher-IQ individuals. Accordingly, measures of cognitive ability predict health outcomes even *within* the same socioeconomic groups, and individuals with higher intelligence seem to make more efficient and better use of the resources that are made available to prevent and improve health problems.

The most impressive source of evidence derives from an almost nationwide study in Scotland, where measures of IQ obtained during childhood predicted individual differences in mortality (including cancer and cardiovascular illnesses) many decades later, even when socioeconomic factors were taken into

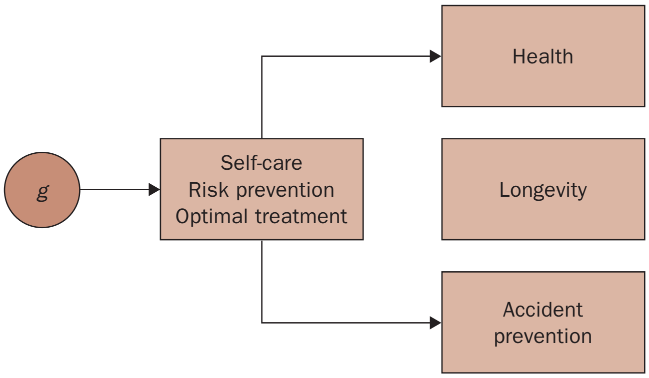

Figure 6.5 Intelligence predicts health and longevity.

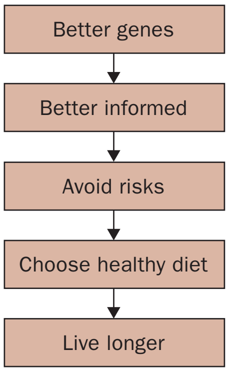

Figure 6.6 The path from *g* to longevity.

account (Gottfredson & Deary, 2004). The Scottish survey examined IQ data for almost every 11-year-old Scottish citizen (N = 87,498) who attended school on June 1, 1932. Many decades later, Ian Deary and colleagues collected archival data and tracked medical as well as death records for thousands of participants. The results (reported in Deary, Whalley, & Starr, 2003) showed that IQ scores at the age of 11 predicted survival rate at the age of 76. Furthermore, participants who scored 1 standard deviation (15 points) lower in IQ had a 27 percent increase in cancer deaths if they were male, and 40 percent if they were female. For stomach and lung cancer deaths, the effects of IQ were found to be even stronger, no doubt due to the socioeconomic factors associated with these forms of cancer.

Similarly, Hart et al. (2003) reported that for every 15-point reduction in IQ scores there was a 17 percent increment in death risk (or 12 percent when socioeconomic factors were partialled out). The impact of cognitive ability on longevity is stronger in deprived or poor social groups, indicating that *g* moderates the correlation between socioeconomic status and mortality. Likewise, higher socioeconomic status may moderate the impact of IQ on longevity. Hart et al. (2003) also found that childhood IQ scores predicted the likelihood of dying from heart-related diseases and lung cancer. This is consistent with Gottfredson's (2004a) findings and the idea that brighter individuals are more likely to choose healthier diets and avoid or give up smoking.

Accordingly, cognitive ability seems to have pervasive effects on health outcomes. Whilst there may be several mediating and moderating variables underlying the correlation between IQ and health factors, the mechanisms by which cognitive ability may lead to positive health are essentially no different than the ones affecting job or academic performance. Higher intelligence provides individuals with faster, better, and more efficient reasoning and learning ability and when this ability is applied to understanding the causes of good and bad health, IQ is no doubt advantageous for health-related decision-making. Thus reviewers argued that:

> Dealing with the novel, ever-changing, and complex is what health self-care demands. Preventive information proliferates, and new treatments often require regular self-monitoring and complicated self-medication. Good health depends as much on preventing as on ameliorating illness, injury, and disability. Preventing some aspects of chronic disease is arguably no less cognitive a process than preventing accidents, the fourth leading cause of death in the United States, behind cancer, heart disease, and stroke. (Gottfredson & Deary, 2004, p. 2)

One advantage of cognitive ability tests over personality and other latent psychological constructs that may be tested as predictors of health outcomes is that they provide an objective measure of individual differences. Thus they are not exposed to socially desirable responding or faking, as are self-reports. However, personality traits have also been shown to affect health outcomes (as seen in section 3.6). In that sense, it would be important for studies to examine the joint impact of personality and intelligence factors on health-related behaviors as well as longevity. This would provide an indicator of the extent to which individuals' health may depend on their dispositions, preferences, interests, and abilities. For example, some individuals may choose to indulge in risky behaviors whilst being aware of the consequences of their acts, whereas others may avoid such behaviors without necessarily knowing that they have made the "right" choice. Likewise, health-advantageous and disadvantageous individual difference factors may combine in the same individuals, such that a person may score high on IQ as well as on Psychoticism or sensation-seeking. Understanding such interactions would no doubt enhance our understanding of individual differences underlying health outcomes.

6.6 INTELLIGENCE AND SOCIAL CLASS

The idea that social class may in part be a consequence of individual differences in cognitive ability (rather than its cause) is no doubt controversial and has important political and sociological implications. Yet, evidence that:

a) individual differences in intelligence *precede* and are more stable over time than socioeconomic status, and that
b) both constructs are highly *intercorrelated*

has led mainstream intelligence research to emphasize the importance of both acknowledging and understanding the social consequences of *g* (see also Box 6.1).

Box 6.1

The Bell Curve controversy

Although genetic and group differences in IQ have concerned differential psychologists for over a century, much of the controversy surrounding these and related themes was sparked by the publication of *The Bell Curve* (Herrnstein & Murray, 1994), a bestselling book that takes its name from the normal distribution of IQ scores (see again section 5.3.3 and Figure 5.3). In it, the authors assess the impact of intelligence in the United States, including a wide range of social, economic, and political consequences of differences in cognitive ability.

Fundamentally, Herrnstein and Murray argue that increases in socioeconomic status are the consequences of a "cognitive elite," i.e., a group of individuals with a higher IQ. Thus success in life is not due to socioeconomic advantages but is the result merely of higher levels of cognitive ability. Accordingly, and most controversially, social deprivation is not a cause of lower IQ scores but its very consequence. Such claims led critics to accuse the authors of "scientific racism."

Whereas the book reports numerous statistics (particularly significant correlates of IQ), the strength of its argument relies on the *heritability* estimates for cognitive ability – in the range of 40 to 80 percent (see chapter 7) – that is, the fact that there are strong *genetic* influences on intelligence, notably psychometric *g*. Whereas this idea is not new (it had been anticipated by Arthur Jensen in the late 1960s), Herrnstein and Murray linked this argument with socioeconomic factors in an unprecedented manner. Furthermore, the authors "praise" the US economy as a model of meritocracy and highlight the importance of a society where wealth is distributed on the basis of intelligence rather than social class. However, this also implies that disadvantaged or unsuccessful individuals are responsible for their own misfortune and that little can be done to reverse inequalities.

The authors base their case on longitudinal evidence derived from analyzing archival data on the National Longitudinal Study of Youth. Information was available on the Armed Forces Qualifying Test (a sort of IQ test) and subsequent socioeconomic variables. Herrnstein and Murray found IQ scores to be a better predictor than socioeconomic status of most socioeconomic outcomes than was social class background. In fact, after partialling out IQ scores, several race differences in socioeconomic outcomes seemed elusive.

Critics such as Leon Kamin regarded the book as "a disservice to and abuse of science," whereas Thomas Sowell criticized the authors for drawing partial conclusions in order to hold their argument. On the other hand, Jared Diamond argued that group differences in socioeconomic status are a result of geographical factors like terrain and natural resources. Yet *The Bell Curve* contains relatively moderate, and mostly implicit, views on the implications of genetic differences in cognitive ability, and even some of its critics have considered it a thorough and honest proposition.

More importantly, differential psychologists have been quite unanimous in their support for *The Bell Curve*. In fact, in the year the book was published, 52 eminent intelligence experts (not only from differential psychology) published a dossier entitled "Mainstream Science on Intelligence" in which they endorsed the core claims and data presented by Herrnstein and Murray.

Socioeconomic differences in cognitive ability are by no means a new finding. Many decades ago, Terman estimated a 14-point IQ gap between the "lowest-" and "highest-"class children tested. Likewise, early demographic studies found that suburban samples scored lower on IQ tests than urban ones, though this finding was attributed to the fact that "brighter" people were more likely to migrate to cities (Cattell, 1937; Terman & Merrill, 1937; Thomson, 1921). Whilst ability differences between urban and suburban samples have tended to disappear over time, this may simply be a matter of migration and educational changes in demography, particularly in developed and industrialized countries.

There is evidence for both genetic and non-genetic or environmental causes of socioeconomic differences in IQ. For example, adoptive studies suggested that being adopted into a wealthier family tends to increase children's IQ (Mackintosh, 1998). However – as will be seen in chapter 7 – genetic influences seem stronger and more pervasive than environmental ones, meaning children tend to resemble their biological parents (in both personality and intelligence) more than their adoptive ones.

The average correlation between social class and IQ is approximately .55 and seems to persist generation after generation (Jencks, 1972; McCall, 1977). On the other hand, fathers and sons tend to differ more markedly in socioeconomic status (approximate $r = .35$) than in IQ (approximate $r = .50$). Crucially, generational decreases in IQ tend to be associated with decreases in socioeconomic status, whereas generational increases in IQ tend to be associated with increases in socioeconomic status (see Figure 6.7).

Thus Mackintosh (1998, p. 147) argued that "it is difficult to resist the conclusion that such an effect is partly responsible for the maintenance of the correlation between IQ and social class in each generation, and therefore that the direction of causality is partly that IQ differences cause social-class differences rather than simply imply that social-class differences cause IQ differences."

Few studies, however, simultaneously examined the direct effects of both environmental and genetic factors on the IQ–social status correlation. In one of the rare exceptions to this rule, Capron and Duyme (1989) managed to find four groups of

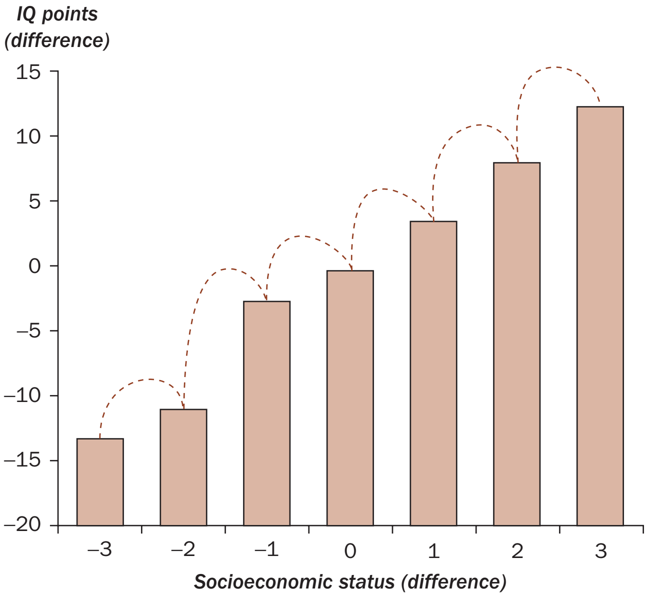

Figure 6.7 Generational gains (father–son) in socioeconomic status as a function of IQ.
Source: Based on Mascie-Taylor & Gibson (1978).

Table 6.4 Adopted children's IQ scores as a function of their biological and adoptive parents' social status

	Adoptive parents	
Biological parents	High social status	Low social status
High social status	119.6	107.5
Low social status	103.6	92.4

Source: Based on Capron & Duyme (1989).

children, two that were born to high social class parents, and two that were born to low social class parents. In turn, each of these groups could be further divided on the basis of whether they were adopted by a low or high social class family. This 2 × 2 (high vs. low status of adoptive × high vs. low status of biological parents) factorial design showed that social class of both adoptive and biological parents had similar effects on the child's IQ, though the sample size was small (i.e., each cell contained only 8–10 participants) (see Table 6.4). Thus IQ scores were highest in children born to high social class parents and adopted by high social class parents (119.6), and lowest in children born to low social class parents and adopted by low social class parents (92.4).

Even if cognitive ability *causes* differences in socioeconomic status, it is important to bear in mind that:

- The correlation between socioeconomic status and IQ is not perfect and, at best, refers to an overlap of 30 percent between most measures.
- Even if this correlation is attenuated and corrected for reliability, there is still a considerable amount of unaccounted variance in socioeconomic status.
- Although there is certainly a *general* tendency for people in one socioeconomic group to obtain a particular type of IQ scores, the rule does not apply to everybody.
- This tendency has implications for the "relative" rather than "absolute" number of individuals from *x* social class that can be found amongst *y* IQ scorers. For example, the number of working-class people tends to exceed, by far, the number of upper-class people, meaning there will be more working-class than upper-class people across most IQ score ranges.

6.7 RACE AND SEX DIFFERENCES IN IQ: FACTS, CONTROVERSIES, AND IMPLICATIONS

No other topic in psychology has been as controversial as the issue of race differences in IQ, in particular the finding that whites tend to have higher IQ scores than blacks. From the early 1920s up to the present day, studies have reported consistent differences of about 10–20 points between the IQ scores of black and white individuals, in favor of the latter (Mackintosh, 1998). This is a robust finding and has been replicated in many countries, though most studies examined US and UK data. Thus Mackintosh (1998, pp. 148–9) concluded: "There can be no serious doubt that North American blacks have an average IQ score some 15 points below that of whites. This difference showed up in the early US Army data, was repeatedly confirmed in subsequent studies between the wars (Shuey, 1966), and has been maintained after the Second War (Loehlin et al., 1975)."

The fact that there are group differences in IQ test performance is a logical and arithmetic consequence of *individual* differences in such tests, one that applies to any measured variable: if some people run faster than others, certain groups will run faster, too; if some people are taller than others, certain groups will be taller, too; if some people get higher IQ scores than others, certain groups will get higher IQ scores, too. It is not the differences but the *causes* and *implications* of such differences that ought to be assessed. However, lay and media reactions have generally preferred to distrust or deny these data, which is a common way of dealing with unpleasant news. It is therefore unsurprising that the media's position on any psychological study showing race or sex differences in IQ is one of skepticism and suspicion. Thus journalists have often questioned the reputation of psychologists reporting such differences, insinuating a hidden political agenda or accusing them of right-wing activism.

Such insinuations or accusations, however, would also imply that there is compelling evidence against the idea of sex or race differences in IQ, when such evidence may be elusive. As a matter of fact, this is misleading and reflects a lack of understanding of the processes underlying sound scientific research. The quality of scientific investigations is judged by experts on the basis of methodological, empirical, and theoretical rather than political grounds. As Mackintosh (1998, p. 149) has argued, "however suspect the motives of many of those who use these data, and however strongly one may deplore their political aims, it is

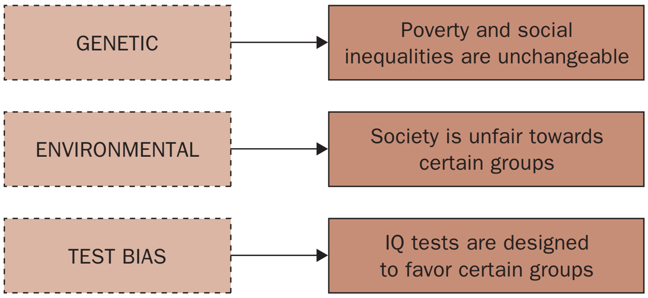

Figure 6.8 Origin of race differences in IQ and implications.

questionable to suppose that much will be gained by pretending that the data do not exist or by refusing to discuss them at all."

Ironically, media attempts to politicize the issue of sex or race differences in IQ have overemphasized the importance of such findings. Furthermore, the media's attempt to deny these differences introduces and consolidates the idea that such differences exist, increasing, reaffirming, and perpetuating the "war of the sexes" and racial prejudice when it supposedly wants to avoid it. The question, in short, is not whether group differences in IQ exist but how significant they are, that is, whether they help us explain real behavioral outcomes or not.

There can be no doubt that interpreting the effects of group differences in IQ (and, in fact, any individual difference variable) is not straightforward. This is where the real debate takes place, as psychologists have long been divided on the basis of whether they ignore, emphasize, or deny the importance or consequences of group differences in IQ. There have been three major theoretical positions when it comes to interpreting such differences, namely, attributing them to genetic, environmental, or measurement factors (see Figure 6.8). This applies not only to race differences but also to other group differences in IQ. Each of these positions has intrinsic consequences and implications for social policy.

Attributing IQ differences to genetic factors seems to imply that there is little to be done to reverse the social inequalities between different groups, and that certain people are just naturally disadvantaged to compete for resources and do well in life. On the other hand, the opposite side of the argument is that environmental factors underlie the causes of group differences, such that socioeconomic status is the real cause of differences in cognitive ability. Thus changing the rules and increasing social justice and educational resources for deprived individuals may eventually lead to IQ gains. Last but not least, the argument that such differences are a mere artifact of psychometric tests – such that, say, the choice of questions or problems is unfair towards certain groups of individuals – would oppose the use of IQ tests in applied or educational settings. Indeed, this argument posits that group differences are artificially created by test designers to justify excluding or favoring certain individuals or groups.

The fact is that even the most controversial reports on race differences in IQ have been careful to interpret the causes of these differences (see again Box 6.1). Indeed, a recent issue of *Psychology, Public Policy, and Law* was entirely devoted to the topic of race differences in IQ. The most radical piece in this issue, written by Rushton and Jensen (2005) – two longstanding advocates of genetic and race differences in IQ – concluded that race differences in intelligence are mainly due to brain size and reflect 80 percent genetic to 20 percent environmental influences.

6.8 SEX DIFFERENCES IN IQ

Sex differences in IQ have sparked almost as much debate as race differences, no doubt due to their sociopolitical implications. Though not always admitted, there is some evidence for the fact that men tend to have an advantage over women on full-scale IQ tests. Most intelligence researchers accept, rather diplomatically, that men do better at some (spatial/mathematical) ability tests, whereas women do better at other (verbal) ability tests. Thus the choice of test may partly determine whether there are gender differences or not. Indeed, some have argued that IQ tests are specifically designed to cancel out rather than reflect differences in intelligence, meaning overinclusion of female-friendly ability problems. Thus Evans and Waites (1981, p. 168) argued that "the two sexes were *defined* to have equal intelligence rather than *discovered* to have equal intelligence" (emphasis in original) (see also Garcia, 1981; Rose, Kamin, & Lewontin, 1984). On the other hand, implicit political censorship is likely to intimidate researchers who believe in sex differences and encourage those who deny them. The most notorious example is Chris Brand, who lost his academic position at the University of Edinburgh after publishing a book containing explicit views on group differences in IQ.

However, almost all pioneers in intelligence testing (e.g., Binet, Burt, Terman) believed there were no sex differences in cognitive ability, and few would claim that these researchers were concerned at that time with "balancing" items to cancel out sex differences (see Mackintosh, 1998, for a review). Thus Terman (1916, pp. 67–70) concluded that "when the IQs of the boys and girls were treated separately there was found a small but fairly constant superiority of the girls up to the age of 13 years, at 14 however the curve for the girls dropped below that of boys . . . however the superiority of girls over boys is so slight . . . that for practical purposes it would seem negligible."

Subsequent reports yield somewhat ambiguous evidence. For example, Wechsler (1944, p. 106) admitted that in standardizing his IQ scale he had taken out items that were probably biased against women. Yet, he also argued that "we have more than a 'sneaking suspicion' that the female of the species is not only more deadly but also more intelligent than the male" (p. 107). Indeed, women seemed to outperform men on early versions of Wechsler's scales as well as the original Stanford/Binet scale.

In recent years, differential researchers – notably Richard Lynn – have launched a systematic series of studies into sex differences in IQ (see also Jensen & Reynolds, 1983; Reynolds, Chastain, Kaufman, & McClean, 1987). Accordingly, there are differences in favor of men rather than women in the region of 3 to 5 points, though probably larger in IQ batteries that include spatial ability tests (Lynn, 1994). Indeed, there are no doubts about men's superior spatial intelligence, a finding that ties in with psychometric as well as biological evidence as spatial ability is related to testosterone levels. Thus females overexposed to androgens (male sex hormones) tend to obtain significantly higher spatial intelligence scores than control groups, albeit not differing in overall IQ (Resnick, Berenbaum, Gottesman, & Bouchard, 1986). Some have argued that sex differences in spatial intelligence are indicative of different evolutionary sex roles, particularly men's past as hunters or gatherers. Alternative (yet compatible) explanations point towards sex differences in *lateralization*, such that spatial abilities are more dependent on right hemispheric activity, though such claims remain contested. On the other hand, there is some evidence for the idea that women slightly outperform men on verbal ability tests (Feingold, 1988; Hyde & Linn, 1988; Mackintosh, 1998). Last but not least, when intelligence is measured in terms of non-verbal reasoning, such as through Raven's Matrices, results have sometimes shown female superiority, sometimes male, and sometimes no significant differences at all (Court, 1983).

Another famous argument (in favor of male intellectual superiority) is that the *distribution* of IQ points is different in women and men. Thus men are more often found amongst both lowest (below 70) and highest (above 140) scorers, whereas women tend to be more homogeneous and less frequently obtain very low or very high scores. In a similar vein, Lubinski and Humphreys (1990) looked at the sex distributions in IQ in a sample of approximately 100,000 teenagers and found the standard deviation for males to be 7 percent larger than for females. This led Lynn (1994) to the controversial assumption that IQ is the reason for the unequal sex ratio amongst eminent figures in the arts, sciences, and politics, who tend to be male rather than female, an assumption that is obviously controversial because it undermines a wide range of socioeconomic and political factors that have historically disadvantaged women in relation to men.

It has also been hypothesized that men's higher IQ score may be a direct consequence of their larger brain sizes, a claim that has been backed up by consistent evidence of correlations in the region of .30 (though usually lower) between brain size and IQ scores (Rushton & Ackney, 1996). Controversially, correlations between brain size and IQ have also been used to support the idea of race differences in IQ, as whites tend to have larger brains than blacks. Yet, these two assertions may be incompatible as (1) the average brain volume differences between blacks and whites is at least five times smaller than that between men and women, and (2) the average IQ difference between men and women is only one-third of the average difference between whites and blacks. Furthermore, Asian groups (e.g., Chinese, Japanese) tend to score significantly higher on IQ tests than whites despite not having significantly larger brains.

If publishing data showing gender differences in intelligence has been controversial and arguing otherwise may increase academics' popularity, attacking the concept of IQ has made many experts rich and famous. Ever since primary school exams, the mere idea of being tested or examined is bound to evoke anxiety and fear of failure. It is therefore unsurprising that few individuals enjoy being tested and put under pressure by a psychometric test that may not only tell them how bright they are, but also decide on their future career. This may also explain the media and commercial success of the current crusade, led by the likes of Sternberg, Gardner, and Goleman (see chapter 8), to destroy the reputation of IQ tests and attempt to replace the traditional notion of intelligence with other, more fashionable and "user-friendly" abilities, such as social, practical, and, in particular, emotional intelligence.

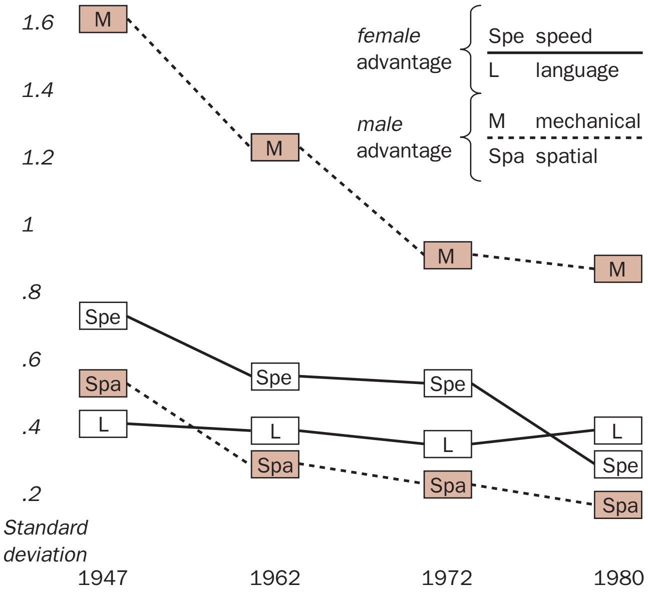

Figure 6.9 In decline: sex differences in abilities.
Source: Based on Feingold (1988).

Although these "abilities" (which, by the way, are supposed to be higher in women than men) have met with wide lay enthusiasm, most academics remain unconvinced as to whether they provide any reliable, novel, or useful information and there are three major reasons underlying this skepticism:

1. It is not possible to design objective tests of emotional, practical, or social intelligence, which means these abilities can only be assessed through self-report inventories.
2. When assessed through self-reports, most novel abilities are substantially correlated with established personality traits, such that constructs like emotional or social intelligence may only be new names for known individual difference constructs (for instance, emotional intelligence may simply refer to a combination of low Neuroticism and high Extraversion).
3. Most novel intelligence theories are largely based on the assumption that traditional IQ tests are not a valid indicator of a person's real intelligence.

6.9 EVEN MORE BASIC: DECOMPOSING INTELLIGENCE

A number of researchers have also looked at "lower" correlates of *g*, with the idea of identifying the causes rather than consequences of intellectual ability. These approaches have aimed at pinpointing the very basic component of *g* in the hope of obtaining a more biological, less cultural, "rawer" measure of brain efficiency. Intelligence researchers have for many decades speculated on the possibility of *g* being ultimately a measure of neural efficiency or neural speed (Anderson, 1992; Eysenck, 1982; Jensen, 1998; Spearman, 1904). In fact, as seen in chapter 2, elementary cognitive processes represented an essential aspect in early experimental approaches to intelligence. The idea underlying these approaches is simple: more efficient brains should be capable of faster and more accurate processing, which in turn is advantageous for information acquisition. Thus "the 'intelligent' nervous system will respond accurately to incoming signals, and will therefore also be able to respond rapidly; the less intelligent will make errors and respond slowly" (Mackintosh, 1998, p. 233). Within this paradigm, two types of task have received widespread attention:

reaction time a measure of the speed of intellectual processing in which a stimulus (e.g., a light) is seen until a decision is made by the participant and a response enacted

inspection time a measure of the speed of intellectual processing in which a stimulus (e.g., lines of different lengths) is presented and inspected for a very short time before being removed

a) ***Reaction time***: this simply requires participants to "react" to a signal (sound or visual stimuli) by pressing a key; alternatively, *choice* reaction time experiments combine different signals which participants need to discriminate before reacting.
b) ***Inspection time***: this requires participants to "inspect" characteristics of perceptual stimuli, such as comparing the length of two lines flashed briefly. The experimenter manipulates the time of exposure to affect individuals' response and error rate.

event-related potential a brain response to an internal or external stimulus, measured by a procedure known as electroencephalography (EEG) which measures electrical activity of the brain through electrodes placed on the scalp

Both reaction and inspection time performance have been consistently correlated with measures of *g* and IQ. Furthermore, studies have also explored the relationship between these simple information processing measures, IQ, and measures of brain functioning such as **event-related potentials** (ERPs) (see Figure 6.10). Correlations between ERP and IQ tests led Eysenck (1982, p. 6) to suggest that "we have come quite close to the physiological measurement of the genotype underlying the phenotypic IQ test results on which we have had to rely so far" (see also section 6.7 on race differences in IQ). Yet, measures of reaction time, information processing, perceptual speed, or inspection time as measures of neurophysiological activity are poorer predictors of learning ability and educational/occupational outcomes than are cognitive ability tests. In fact, recent findings by Ackerman and Heggestad (1997) show that crystallized abilities are more useful predictors of intellectual performance outcomes than are fluid or more "biological" markers of *g* (see also Chamorro-Premuzic & Furnham, 2005). For example, Rushton and Ackney (1996) report correlations between brain size and measures of academic and job performance in the region of .38. Yet, psychometric measures of *g* and IQ tend to have higher validities (as seen in sections 6.3 and 6.4).

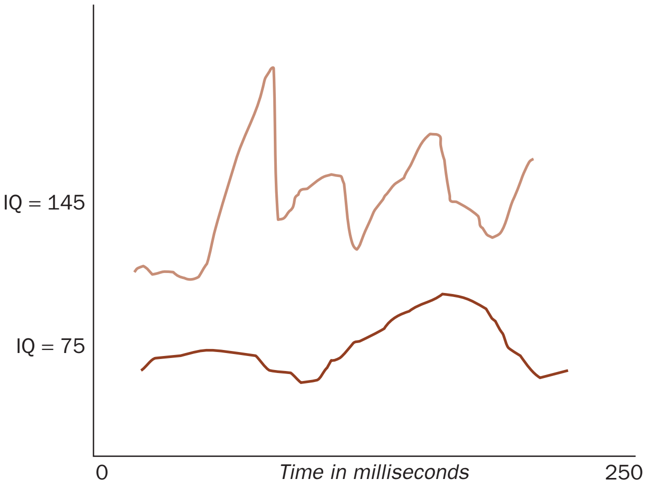

Figure 6.10 Event-related potentials for low and high IQ subjects.
Source: Based loosely on Ertl & Schafer (1969).

On the other hand, studies have examined measures of brain activity via electroencephalograph (EEG) records. EEG waves signal changes in mental states, for example, engaged, drowsy, asleep, and there are clear individual differences in such patterns. The question, however, is whether such differences have any important relation with measures of cognitive ability (i.e., psychometric tests) and, if they do, what they mean. For example, studies have found that the difference in brain activation between states of rest and cognitive task performance is less marked in individuals with higher than lower IQs (Giannitrapani, 1985). This is consistent with the idea that individuals with higher IQs use their brain more efficiently and "tend to have a relatively lower rate of energy use (as measured by glucose metabolism)" (Gottfredson, 2004b, p. 38).

This does not tell us about the causes of cognitive ability but may, on the contrary, reflect the fact that higher intelligence may lead to reduced "brain consumption," to put it metaphorically. In fact, this type of interpretation applies to most correlational studies between brain (physiological) and behavioral (psychometric) outcomes, whereby cognitive ability may simply be a nexus or mediator between the two measures, that is, influences both brain activity and cognitive performance (see Figure 6.11).

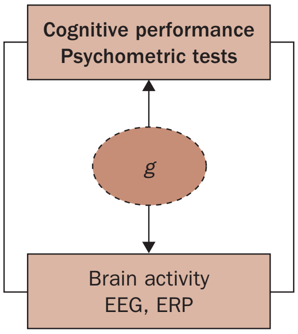

Figure 6.11 *g* may influence both cognitive performance and brain activity measures.

6.10 SUMMARY AND CONCLUSIONS

This chapter was concerned with the validity of intelligence, that is, the question of whether intelligence tests are useful predictors of real-life outcomes. In addition, this chapter reviewed some of the literature on the causes of group differences in IQ. As seen:

1. There is little doubt today among academic psychologists that good IQ tests represent an excellent indicator of an individual's potential for achievement in the real world, in particular when adaptation to novel, complex environments is required.
2. However, IQ tests are hardly the only indicator of an individual's ability to succeed in life. Even academic performance, which has been the validity criterion for IQ tests for more than a century, is dependent on factors other than IQ. Thus intelligence is necessary but not sufficient. Whereas a high IQ will never be a drawback per se, intermediate or low IQ levels will not necessarily preclude individuals from performing well on most everyday tasks. This will be the case even in the long run, provided they are able and willing to compensate with other aspects of their personality, such as being stable, confident, motivated, organized, or hardworking. Likewise if people lack confidence, stability, and motivation, and are unwilling to work hard, IQ scores will be a poor predictor of performance.
3. Claims that gender differences in IQ are responsible for the achievement gap between women and men are exaggerated and show an incomplete picture of the multiple determinants of individual differences in achievement. Furthermore, failure to account for gender differences in self-assessed abilities, vocational interests, and motivational factors exposes the limitations of traditional ability measures, producing incongruent interpretations of findings. Until the combination of factors determining educational and occupational success and failure is fully understood, the implications of the possible gender gap in IQ will remain a matter of political speculation rather than scientific evidence.
4. Ever since the 1960s, the idea that individual differences in intelligence, as measured by IQ tests, may have a strong genetic or hereditary component has been at the center of a heated academic and political debate. Although mainstream IQ researchers today are in agreement about the strong genetic basis of intelligence, the political implications of such findings are hard to digest, particularly as IQ tests were initially employed to enhance meritocratic selection and facilitate (rather than obstruct) social mobility.

In chapter 7, I examine studies on behavior genetics, which attempt to estimate the degree to which genetic and environmental factors influence personality and intelligence.

KEY READINGS

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*, 219–245.

Brody, N. (2000). History of theories and measurements of intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 16–33). New York: Cambridge University Press.

Chamorro-Premuzic, T., & Furnham, A. (2006). Intellectual competence and the intelligent personality: A third way in differential psychology. *Review of General Psychology*. In press.

Deary, I. J. (2001). *Intelligence: A very short introduction*. Oxford: Oxford University Press.

Gottfredson, L. S. (2004). Intelligence: Is it the epidemiologists' elusive "fundamental cause" of social class inequalities in health? *Journal of Personality and Social Psychology*, *86*, 174–199.