GALTON'S LEGACY TO RESEARCH ON INTELLIGENCE

ARTHUR R. JENSEN

Journal of Biosocial Science / Volume 34 / Issue 02 / April 2002, pp 145 - 172
DOI: 10.1017/S0021932002001451, Published online: 28 March 2002

J. biosoc. Sci. (2002) **34**, 145–172

GALTON'S LEGACY TO RESEARCH ON INTELLIGENCE

ARTHUR R. JENSEN

School of Education, University of California, Berkeley, USA

Summary. In the 1999 Galton Lecture for the annual conference of The Galton Institute, the author summarizes the main elements of Galton's ideas about human mental ability and the research paradigm they generated, including the concept of 'general' mental ability, its hereditary component, its physical basis, racial differences, and methods for measuring individual differences in general ability. Although the conclusions Galton drew from his empirical studies were seldom compelling for lack of the needed technology and methods of statistical inference in his day, contemporary research has generally borne out most of Galton's original and largely intuitive ideas, which still inspire mainstream scientific research on intelligence.

Introduction

Consideration of Sir Francis Galton's legacy to research on human intelligence immediately requires making an important distinction between the two major aspects of his work in this domain. The first aspect was his extraordinary curiosity, inventiveness and investigative zeal. These generated astute hypotheses, fertile methodologies, and empirical discoveries of historical scientific importance, which laid the foundations for psychometrics and human behavioural genetics. The second aspect was his promotion of eugenics as a secular religion aimed at the future well-being of humanity. It seemed obvious and even unarguable to Galton that, from a eugenic viewpoint, superior mental and behavioural capacities, as well as physical health, are advantageous, not only to an individual but for the well-being of society as a whole.

In his later years Galton's pioneer contributions to the science of mental ability became so amalgamated with his enthusiasm for eugenics as to have also contributed to the disfavour in which Galtonian research on human intelligence has been held in the latter half of the twentieth century. Even more damaging has been a systematic inculcation in the public mind of a link between the humane concept Galton termed 'eugenics' and the fanatical political abuses and horrific acts committed in its name which were revealed at the conclusion of World War II. This has probably contributed to stigmatizing research on intelligence along Galtonian lines in a manner seldom seen for scientific research on any other subject, save perhaps for evolution as perceived by Biblical Fundamentalists.

What seems not to have been sufficiently clear, even to Galton himself, yet needs to be emphasized, is this: prescriptive eugenics falls not in the purview of science, but in the province of moral philosophy. Despite the humanitarian aims espoused by Galton, the desirability of eugenic policies is entirely an ethical question, rather than a scientific one. Science deals strictly with what *is*, not with what anyone thinks *ought to be*. This critical distinction in no way belittles the eugenics concept or suggests that its aims, means or feasibility be divorced from scientific scrutiny. The very same distinction must be made between medical science and medical practice, especially as biomedical technology advances.

The purpose here is to outline Galton's contribution to the scientific study of mental ability and assess its impact in the light of present-day research.

The Galton paradigm

Among his many diverse scientific accomplishments, Galton is also recognized as the founder of 'differential psychology', the study of individual and group differences in psychological traits. Galton's main works in this field, from which this review is mostly drawn, are *Hereditary Genius* (1869) and *Inquiries into Human Faculty* (1883).

The Galton paradigm with respect to mental ability can be stated in the form of four research questions, to each of which Galton turned his attention and proposed at least tentative answers. They are: (1) the idea of 'general mental ability'; (2) the 'inheritance' of individual differences in this general ability; (3) the 'measurement' of general ability; and (4) 'racial differences' in general ability.

General mental ability

In his preface to the second edition of *Hereditary Genius* (1892), Galton expressed regret that he had not titled his book *Hereditary Ability*, as his main interest was not in genius *per se*, but in the causes of variation in human abilities. In his day, without suitable mental tests, he could study the genetic inheritance of ability only by examining the pedigrees of eminent individuals who had displayed unquestionably outstanding intellectual achievements.

Galton usually referred to the mental qualities that made outstanding achievements possible as 'natural ability'. Not only did he conceive of natural ability as largely attributable to heredity, but there seems little question that he regarded it as a broad or 'general' ability. This is indicated in Galton's frequently quoted statement in *Hereditary Genius* (1892):

> 'In statesmanship, generalship, literature, science, poetry, art, just the same enormous differences are found between man and man; and numerous instances recorded in this book, will show in how small degree, eminence, either in these or any other class of intellectual powers, can be considered as due to purely special powers. They are rather to be considered in those instances as the result of concentrated efforts, made by men who are widely gifted. People lay too much stress on apparent specialities, thinking over-rashly that, because a man is devoted to some particular pursuit, he could not possibly have succeeded in anything else. They might just as well say that, because a youth had fallen desperately in love with a brunette, he could not possibly

have fallen in love with a blonde. He may or may not have more natural liking for the former type of beauty than the latter, but it is more probable as not the affair was mainly or wholly due to a general amorousness of disposition,' (p. 64).

Because Galton found that the lineages of many eminent persons also contained other individuals of distinction, but often in quite different fields of intellectual endeavour, Galton, in his *Inquiries in Human Faculty* (1883), recognized what he specifically termed the 'convertibility' of the hereditary component of variation in mental ability (p. 57), meaning that the hereditary component of 'natural ability' is, at least in part, a 'general cognitive ability' that could be manifested in many different kinds of intellectual achievement. The view that variation in a natural, general cognitive ability constitutes the main source of individual and group differences in intellectual achievement, both in school and in occupations, therefore, is clear in Galton's writings. But he also recognized 'special abilities' or talents as being necessary but not sufficient for outstanding performance in particular fields, such as music, art and mathematics. He also held that a sufficiently high level of general ability was required for the distinguished expression of any special talent. His study of the genealogies of eminent persons in many fields led him to believe that general ability is more heritable than are special abilities, which are more influenced by environmental opportunity and education and are developed to a distinguished level only through assiduous practice.

It must be noted that Galton never clearly spelt out the psychometric implications of his notion of general ability, such as the conception we owe to Charles Spearman of 'general ability', or psychometric *g*, as the largest common factor, or source of variance, that accounts for the positive correlations among measures of many apparently different abilities. The critical distinction between 'ability in general' as a person's average level of performance on tests of various abilities, and 'general ability' as a common factor, independent of specific abilities, in the positive correlations among individual differences in various abilities, was never explicitly drawn by Galton, implicit though this distinction may seem in some of his writings.

The inheritance of general ability

The term 'heritability' in its technical sense had not yet come into existence in Galton's time. Heritability is the proportion of population variance in a given metric trait (the observed phenotype) that is accounted for by genetic variance in that trait (the inferred genotype). What Galton studied in *Hereditary Genius*, however, was the phenotypical resemblance between eminent individuals and their relatives. From such data he inferred a genetic basis, or inheritance. But 'inheritance' in this sense reflects only that component of *heritability* that relatives have in common. Even between first-degree relatives (parent–child and full siblings) this genetic component is only about 50% of the phenotypic variance for any trait, physical or psychological. The one exception is monozygotic (MZ) or identical twins, with 100% of their genetic factors in common. In *Human Faculty* (1883) Galton was the first to study MZ twins as a means for assessing the influence of genetic factors on behaviour, noting the remarkable degree of resemblance between them as compared with the far lesser resemblance between dizygotic (DZ) twins or between single-born siblings. The twin method introduced by Galton became a principal tool of quantitative genetics and behavioural genetics, and Galton is now recognized as the father of behavioural genetics. From his own studies, he concluded that general mental ability is inherited in the same way and to a similar degree as many quantitative physical traits. Galton made no essential distinction between physical and mental qualities and viewed them biologically as products of the evolutionary process.

The measurement of general ability

Galton was also the father of psychometrics, the measurement of quantitative behavioural traits. To this end he devised a number of methods and apparatuses, and suggested that individual differences in general ability are reflected in performance on relatively simple sensory capacities and in speed of reaction to a stimulus, variables that could be objectively measured by tests of sensory discrimination and reaction time. His attempt to demonstrate this with the data he collected on thousands of people in his anthropometric laboratory appeared unsuccessful at the time.

Galton's so-called 'brass-instrument' approach to measuring intelligence was virtually abandoned in the first decade of the twentieth century, when Alfred Binet's more complex test was adopted by psychologists and educators as a practical measure of intelligence. Galton's instruments and techniques lacked the precision and reliability needed to evaluate his hypothesis properly, and the methods of statistical inference required for interpreting his data, particularly the method known as the analysis of variance, invented by R. A. Fisher in 1923, had not yet been developed. When the analysis of variance was later applied to Galton's data, in 1985, it proved unequivocally that Galton was on the right track (Johnson *et al.*, 1985). Groups that he had classified by educational and occupational level, which were Galton's only criteria for differences in mental ability, showed statistically significant differences in Galton's simple sensory and reaction time tests.

Before the invention of inferential statistics, however, Galton provided psychometrics with some of its most fundamental measurement tools and descriptive statistics, and all of these are still in use. He invented the measures of bivariate correlation and regression (further developed by Karl Pearson), the use of percentile scores for measuring relative standing on various measurements, and the use of the Gaussian or normal curve as a means for scaling variables that theoretically are normally distributed but cannot be directly measured on an interval scale or a true ratio scale (Burt, 1962; Stigler, 1986, Chapter 8). Galton concluded that general ability, or intelligence, is normally distributed in the population, so measures of rank order could be converted to percentile ranks, which in turn could be transformed to normal deviates. Since Galton scaled intelligence in units of the semi-interquartile range of the normal frequency distribution, his scale can be transformed exactly to the scale of IQ used in all present-day IQ tests. Galton's early statistical inventions were soon after further developed or supplemented by mathematical statisticians: partial and multiple correlation and principal components analysis (by K. Pearson), common factor analysis (by C. Spearman) and the analysis of variance (by R. A. Fisher). These statistical innovators acknowledged their debt to Galton.

Group differences

As a thorough Darwinian, Galton (who was Darwin's half-cousin) assumed that racial differences, both physical and mental, were the product of different environmental adaptations in the course of human evolution. This naturally implies a genetic component in the observed racial differences, as genes, mutations and natural selection are the basic mechanisms of evolution. Galton had no means for directly or rigorously testing this hypothesis, but from his extensive field observations during his explorations in Africa, he made an admittedly rough estimate of the average individual's position in the normal frequency distribution of what he conceived as general mental ability, theoretically scaled on 'Anglo-Saxons', and suggested that Africans were 'two grades' below Anglo-Saxons on his scale, a difference equivalent to 1·33 standard deviations or 20 IQ points on present-day IQ tests. This can hardly be called more than a guesstimate, yet remarkably it is close to the median of all studies of Black IQ worldwide (Lynn, 1991).

Sex differences were also considered by Galton, and from his studies of sensory discrimination and reaction time he concluded that men were higher than women in general ability. His evidence for this conclusion was actually too weak to support such a broad generalization, and the best evidence we have today does not support it. Any given collection of various tests, if their scores are merely summed, can show a sex difference going in either direction, depending on the kinds of tests that compose a particular battery. Men on average score higher on certain tests (such as spatial reasoning) than women, who average higher on other tests (such as verbal fluency). Therefore the observed sex difference in any particular test or battery is of little general or scientific interest. Scores based on the summation (or the mean) of any particular collection of diverse tests do not constitute a true scientific construct. Although such composite scores can have large and practically useful validity for predicting certain criterion outcomes, such as scholastic achievement and job performance, they represent only some unique mixture of the various abilities in the composite, which therefore has no clearly defined generality. Such composite scores have been said to measure 'intelligence in general' as contrasted with 'general intelligence'.

Spearman's construct 'general intelligence', or *g*, is conceptually quite different from 'intelligence in general'. It is derived, not from the summation of the scores on various tests, but from all the correlations between them and represents only the variance they have in common, statistically independent of other sources of variance that particular types of tests may have in common (i.e. group factors) and variance that is unique to each test. In other words, *g* is a distillate rather than a mixture. Spearman's (1904; 1927, pp. 70–71) emphasis on this difference between 'intelligence in general' as a mixture of scores on diverse tests and 'general intelligence' as the common factor (or source of individual differences) among all the diverse tests is one of the most important distinctions in contemporary psychometrics since the development of latent trait theory. It has not been sufficiently recognized that Spearman's *g* was the first latent trait to be identified in the history of psychology (Spearman, 1904; Jensen, 2000).

Those who claim a sex difference in intelligence, typically favouring males, have based their evidence exclusively on 'intelligence in general' (e.g. Lynn, 1999). There has not yet been a demonstration of a significant or consistent sex difference in Spearman's *g*, or 'general intelligence', although several methodologically appropriate studies have tried but failed to detect a significant sex difference (Colom *et al.*, 2000; in press; Jensen, 1980, pp. 624–625; 1998a, pp. 536–542). Two methods for testing this hypothesis are crucial; to claim a sex difference in *g* either one or both methods must reject the null hypothesis of no sex difference. First, if there exists a sex difference in *g*, then the *g* loadings of the various tests in a battery should be significantly correlated with the sizes of the standardized mean sex differences on each of those tests. But this has not been found in the several different test batteries in which it has been tried (see studies cited above). Second, if there were a sex difference in *g*, then if we included in the factor analysis of a test battery the point-biserial correlations of sex with each of the tests, the sex variable should have a large loading on the *g* factor. Yet in every study the loadings of sex on the *g* factor have been small, near-zero values not consistently favouring either sex. Sex does have quite large loadings on other factors, however, particularly a spatial reasoning factor (favouring males), but also on a verbal factor (favouring females), and on several smaller factors: verbal fluency, clerical speed and accuracy (favouring females) and quantitative reasoning (favouring males).

The factor analysis of mental abilities

Spearman's technical refinement and transformation of Galton's original but imprecise conception of general mental ability created one of the scientifically most productive constructs in psychology and undoubtedly one of the most challenging, given its relevance and implications for society (Gottfredson, 1997). General mental ability, or *g*, is a major construct in differential psychology. It is important to have a clear conception of what is meant by this term, which now has a precise technical definition.

Before Charles Spearman (1863–1945), who pioneered our present understanding, general ability was conceived to be the sum or average of an individual's scores on a collection of varied mental tests. This measure, arrived at by the simple summation of test scores – now called 'ability in general' to distinguish it from 'general ability' – was Alfred Binet's method of assessing ability. Spearman, however, argued that simple summation of scores on various tests measuring different abilities (for example, vocabulary, arithmetic reasoning, memory span and solving mechanical puzzles) yields just an arbitrary mixture of various capabilities. Without any objective basis for deciding which particular tests should go into the mixture, the composite score therefore is an arbitrary hotchpotch rather than a scientifically interpretable measure (Spearman, 1927, pp. 70–71). According to Spearman, the concept of 'general ability' should be based, not on the *summation* or arithmetic average of a number of different tests, but on the *correlations* among many measures of different abilities. That is to say, general ability is most accurately conceived of as the 'common factor' among a large number of diverse cognitive tests. More fundamentally, it is what causes individual differences in the levels of performance on various tests to be correlated with one another.

Spearman invented a rather complex mathematical algorithm, now called 'factor analysis', for determining the degree to which any given test measures this common factor within the context of a number of other tests. He used this method to examine Galton's idea that simple tests of auditory and visual discrimination measured whatever is common to these sensory tests and to several, more complex, variables based on different types of scholastic achievement and teachers' ratings of pupils' mental ability. Spearman concluded that Galton's hypothesis was correct: all of these diverse measures do, in fact, share a common factor, that is, a common source of variance, or individual differences. Spearman named this general factor simply *g*, for general, because it appeared to be general to all cognitive tests, however diverse they appear by simple inspection.

In the 96 years since Spearman's discovery, *g* has been found to be absolutely ubiquitous, to some degree, in every kind of mental task. All kinds of cognitive performance, however diverse, are found to be positively correlated with one another across the population. This is not a theory, but an empirical fact. Those who, for whatever reason, do not like this fact have never been able to disprove it, even though that would be easy to do if it were actually untrue. Also, it is found that differences in all kinds of mental abilities among people sampled at random from the general population are more attributable to their differences in *g* than to any *special* abilities independent of *g*.

No real theory has yet been developed to explain these remarkable facts. The nascent formulations known as '*g* theory' are in the process of continual discovery and revision. But that is also true of the theory of gravitation, an equally central construct in physics. But aside from theory, the psychometric existence of *g* is one of the best established facts in all of psychology. Similarly, Newton's inverse square law of gravitation is not a theory or an explanation of gravitation, but rather an observed fact *about* gravitation. The same is true of psychometric *g*; it is an empirical fact, but not an explanation of that fact. The little *g* of psychology and the big *G* of physics have much else in common: they both are central concepts in their respective fields of science; their effects are objectively observable and measurable; many important facts and empirical generalizations have been discovered about them and can be predicted from them; there are still various theories and controversies about the causal or explanatory theory of both *g* and *G*. Research seeking scientifically comprehensive theories both of *g* and of *G* are currently in progress.

Following Galton, Spearman at first thought only in terms of a 'general ability' and 'special abilities'. Spearman's earliest formulation, called the 'two-factor theory', posited that each and every kind of mental test reflects a factor (*g*) common to all mental tests and also reflects a factor *s* that is *specific* to a particular test. (Recall that a 'factor', in Spearman's terms, is a latent trait, or hypothetical source of individual differences, measured as a component of variance that any two or more tests have in common.)

However, it was soon discovered that, besides *g* and *s*, certain groups of tests have other sources of variance in common. These Spearman termed 'group factors', because they appear, not in all cognitive tests, but only in certain groups of tests. Some of the established group factors are 'verbal', 'numerical' and 'spatial' abilities.

Other pioneers of factor analysis, including Karl Pearson, Cyril Burt and Leon Thurstone, further developed the mathematical techniques for distinguishing and measuring any number of group factors (F_1, F_2 . . . F_N). (Spearman's simpler method

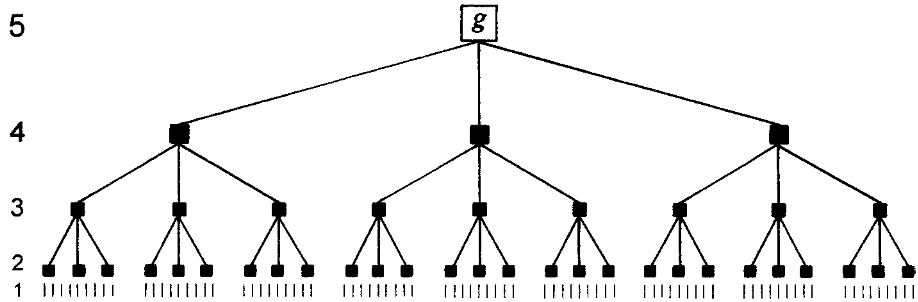

Fig. 1. A hierarchical factor model (for explanation see text).

of factor analysis was not able to extract more than a single factor.) The total variance of the scores on any mental test, then, could be expressed in factor-analytic terms as the sum of the variances attributable to each of a number of independent components of variance:

$$\text{Total variance} = g + F_1 + F_2 + \ldots + F_N + s + e,$$

where s is the test's specificity and e is a random error of measurement.

In light of all the hundreds of factor analyses (Carroll, 1993) that have been performed on various batteries of diverse mental tests since Spearman discovered g in 1904, the psychometric structure of abilities is best represented by a hierarchical model, as shown in Fig. 1. At the very bottom of the hierarchy (Level 1) are single test items, each of which measures some particular capability. Examples are defining the word 'amanuensis', repeating a series of eight digits, or continuing the number series 1, 4, 9, 16, 25, –, –. The items may call upon any particular type of mental capability, provided it can be measured or scored objectively for correctness or quality of response (as in the preceding examples). The number of different kinds of items is theoretically unlimited and constrained only by the inventiveness of test constructors. The next stratum in the hierarchy (Level 2) represents various tests, each composed of a number of items at Level 1 that are more highly intercorrelated with each other than they are with other items at Level 1. There is virtually no limit to the possible number of different tests, either. The 'first-order' factors (Level 3) are derived by factor analysis from all the correlations among the various tests at Level 2. A factor analysis of the correlations between the first-order factors yields a much smaller number of 'second-order' factors at Level 4. Finally, a factor analysis performed on the correlations between the second-order factors yields a single 'third-order' factor, g, at the apex of the hierarchy (Level 5). Thus g can be defined as the highest order common factor in a large battery of diverse cognitive tests. If the number and diversity of the tests (Level 2) is not very large, g emerges at Level 4 as a single second-order factor. It is noteworthy that in the hundreds of factor analyses reported in the literature, g always emerges as either a second-order or a third-order factor (Carroll, 1993).

There are several essential points to be understood about this hierarchy:

- The 'elements' at each successively higher level in the hierarchy represent a greater degree of generality of the ability factor being identified than the elements at the lower levels of the hierarchy. In other words, it is a hierarchy of generality.
- The factors at each higher level in the hierarchy also are more 'distilled', so to speak, than the factors at lower levels, because the variance that the elements at a lower level have in common is moved up to the next higher level. The lower-order factors are thus residualized at each level. In the final result, therefore, each of the factors emerging from the whole analysis is left uncorrelated with each of the other factors. In other words, the hierarchical factor analysis identifies all of the statistically independent (i.e. uncorrelated) sources of variance represented in the collection of tests that have been factor-analysed.
- At each successively higher level in the hierarchy we are further removed from the specific characteristics and informational content of the original tests from which the factors are extracted. Therefore, whatever a given factor at each higher-order represents is more difficult to describe in terms of the information content of the tests or even in terms of the cognitive processes they supposedly called for. First-order factors, being fairly close to the tests from which they are extracted, can usually be described as representing, for example, verbal, or numerical, or spatial, or short-term memory, or other abilities that are recognizable from the particular information content or skills required by those tests with the largest loadings on a given factor.
- Because *g* is at the apex of the hierarchy, it is consequently the most highly 'distilled' of all the factors. It is really impossible, therefore, to describe *g* in terms of any characteristics of the various tests themselves. We are unable to describe *g* even in psychological terms. As Spearman (1927, pp. 75–76) said, we can define *g* by site, but not by nature, meaning that while we can identify the particular tests that reflect *g* the most and the least, we cannot say just what is being reflected. He hypothesized *g* metaphorically as 'mental energy', but he could only note empirically that the most highly *g*-loaded tests are those that involve the 'eduction of relations and correlates', that is, inductive or deductive reasoning. But this does not adequately define *g*, because *g* also appears in tests that do not involve any form of reasoning at all. The nature of *g* itself, therefore, is indescribable in terms or tests or psychological processes. In fact, it has been said that *g* is not really an ability at all. It is a result of whatever causes various abilities to be positively correlated with each other. It is not essentially a psychological variable, but some property (or properties) of the brain that has cognitive manifestations that result in the emergence of *g*. Probably the best we can do in purely psychological terms is to say that *g* reflects the speed and efficiency of information processing. This characterization is essentially Galton's original conception of general mental ability.
- Factor analysis also shows the degree to which each of the tests in the analysis is correlated with each of the independent factors yielded by that analysis. The correlation between a test and a factor is called a 'factor loading'.
- From the results of a factor analysis and, given individuals' test scores, an algorithm can estimate each individual's 'factor score' for each of the factors. Hence, an individual's standing on a given factor (relative to some specified population) can be estimated independently of that individual's standing on *g* or any of the group factors and test specificity.

● The practical predictive validity of mental tests depends largely on their *g* loadings. This is true of tests whether used as predictors of learning ability, educational attainments, occupational level, performance on the job, earned income, health, longevity, or other socially significant variables (Gottfredson, 1997; Jensen, 1998a, Chapter 9). When *g* is statistically removed from test scores, their practical validity and correlations with significant real-life criteria are almost totally lost. It is because IQ is highly *g*-loaded that it is significantly correlated with all these outcomes. In fact, there is a direct relationship between the magnitudes of various tests' *g* loadings and their practical usefulness in schools, colleges, industry and the armed forces. Other factors besides *g* are useful predictors of certain specialized achievements, given an adequate level of *g*, but *g* is still the single best predictor of success for all jobs. Other ability factors and traits of personality used in combination with *g* may yield a somewhat better prediction in some cases. The predominant and ubiquitous correlations of *g* with so many variables of educational, economic and social significance leaves no doubt that it is one of the most important indicators of human capability (Gottfredson, 1997). It is not the only indicator, of course, and to be of value, either to an individual or to society, it must be accompanied by other desirable qualities, among which Galton especially noted energy, industry and persistence.

Now that *g* has been explained as the accepted meaning of 'general intelligence' in modern psychometrics, we can see how Galton's ideas on this subject have fared in recent research.

The heritability of *g*

This subject was the main theme of the 1996 Galton Lecture by Thomas Bouchard (1996). Galton was mostly right: *g* is now certainly known to be highly heritable, although not as heritable as certain polygenic physical features such as stature and fingerprints. Heritability, symbolized as h^2, is defined by the equation:

$$h^2 = (\text{genetic variance})/(\text{phenotypic variance}),$$

and is equivalent to the squared correlation between phenotypes and genotypes. The value of h^2 for a given trait cannot be measured directly but can be estimated in the way Galton originally suggested: from the correlations between relatives with differing degrees of genetic kinship, from monozygotic twins (who have all of their genes in common) to completely unrelated individuals, and with every degree of kinship between these extremes.

The most important findings of recent research on the genetics of mental ability can be listed in a number of points (the specific literature on each of the following points is cited elsewhere: Jensen, 1998a):

● There is no single overall value for h^2, because it gradually increases across the life span, from about 0·40 in early childhood to about 0·80 in later maturity. This increasing expression of genetic tendencies with increasing age goes counter to environmentalist expectations and is evidence for the heritability of IQ.

● The most highly heritable source of variance in different mental tests is the *g* factor. It is the single best predictor of a test's heritability, increasingly as we go from younger to older age groups. The correlation between various tests' *g* loadings and their respective h^2 values is about +0.70. One obvious implication of the high heritability of *g* is that it necessarily has a physical basis in the brain's physiology.

● In children born to parents who are genetically related we see a significant lowering of test scores (about 5 to 7 IQ points for the offspring of cousin matings), an effect known as 'inbreeding depression'. The degree to which various mental tests display inbreeding depression is directly related ($r = +0{\cdot}80$) to their *g* loadings (Jensen, 1983). It is axiomatic that assortative mating for a given heritable trait increases the genetic variance in the population on that trait. More evidence for the pervasiveness and importance of *g* is the fact that the degree of assortative mating varies for different mental tests and is directly related ($r = +0{\cdot}95$) to the tests' *g* loadings (Nagoshi & Johnson, 1986).

● Probably the most far-reaching development in the study of the heritability of *g* has come from the recent applications of the methodology of molecular genetics to this issue. Genes are specific sections of DNA, each having a definite genetic code that drives a particular chemical effect on the growth and function of an organism's physical structure. The specific product of each gene affords a clue to some aspect of the causal chain of physical mechanisms through which the gene affects some end-result. This may be some aspect of behaviour, including behaviour that we recognize as an ability (or a disability, such as dyslexia). Typically, each of these behavioural tendencies is a product of a combination of multiple genes interacting with certain environmental conditions to produce their observable effect.

The techniques of molecular genetics have been adapted to the study of Quantitative Trait Loci (QTL), the search for the loci, in the chromosomes and in the DNA, of the specific genes that contribute to individual differences in some polygenic trait, such as psychometric *g*. A leading pioneer in this research is Professor Robert Plomin, at the Behavioural Genetics Research Unit in the Institute of Psychiatry of the University of London. The genes or QTLs involved in a number of special abilities and disabilities aside from *g* have been studied in this manner (Chorney *et al*., 1998; Plomin, 1997). Because *g* accounts for so much of the genetic variance that special abilities have in common, Plomin is now focusing on the QTL analysis of *g* itself.

The search for the genetic basis of *g* involves administering a highly *g*-loaded psychometric test to large numbers of young people and then comparing sections of DNA from a group who scored in the average range of IQ with the DNA from a group with very high scores. Differences between these groups in the relative frequencies of genes at specific loci in the DNA reveal potential candidates for the genes (or their alleles) that contribute to variance in *g*. The reliable identification of a particular gene requires that the study must be replicated two or three times with new groups that differ in their levels of *g*. Although this work is difficult and time-consuming, we know from the massive research on the heritability of *g* that the genes for *g* are necessarily there; so even if it seems like looking for the proverbial needles in a haystack, with persistence the genes for *g* will be identified. The technology of DNA analysis is rapidly advancing, and the progress of QTL analysis is hastening. So far, Plomin and his co-workers have securely identified three DNA loci that contribute to variance in *g*, and the search continues.

These loci associated with *g*, however, can be truly called genes only when their characteristics in terms of their chemical products and effects have been fully spelt out. The ultimate discovery of the genes for *g* will be not only of scientific interest, but it holds promise of a quantum leap in the understanding and the possibility of favourably influencing mental abilities and disabilities far beyond anything that has ever been achieved by conventional psychological or educational interventions.

Environmental contributions to individual differences

The proportion of variance in *g* attributed to environmental, or non-genetic, influences steadily decreases from infancy to late adolescence. However, it is important to distinguish between two sources of environmental effects, now termed the 'shared' environment (i.e. non-genetic family influences that make children reared together more similar than children who are reared in different families) and the 'non-shared' environment (i.e. non-genetic effects that cause children reared together to differ from one another; Rowe, 1994).

The most surprising discovery is that the *shared* proportion of the total environmental variance in *g* decreases to virtually zero by late adolescence. Early family influences associated with social class, parental education, income and family differences in child-rearing practices common to siblings reared together have little or no effect on an individual's intelligence as an adult, provided that the individual has grown up within the normal range of humane social environments. This has been proved consistently in studies of adopted individuals, whose IQs prove to be correlated with the IQ and educational level of their adoptive parents in early childhood but by late adolescence are more highly correlated with the IQ of their biological parents from whom they were separated as infants.

The variance due to non-shared environmental influences remains fairly constant throughout the life span. Much of it is probably attributable to what is termed the 'microenvironment'. This includes countless small events, mostly random and accidental, each of which has some small effect, for good or ill, on mental development. These are mainly physical effects related to health and nutrition that occur prenatally and in infancy that affect mental development. The argument that these microenvironmental influences largely constitute the non-shared and non-genetic variance in IQ has been spelt out elsewhere (Jensen, 1997a).

Probably one of the reasons that intentional educational and psychological manipulations have shown so little effect on the development of *g*, despite the large proportion of non-genetic variance in *g*, is because much of this non-genetic variance results from innumerable, small, random, physical effects on brain development. Many such effects are hardly amenable to being brought under systematic environmental control. It seems likely that even the purely environmental aspect of *g* is caused more by what could be termed 'biological noise' in the course of the individual's development than by the purely social-psychological influences in the normal range of environments. Although much still remains to be learnt about the nature of the *non-shared* environmental influence, a fuller understanding probably offers more promise for influencing mental development favourably than does manipulation of the variables that constitute the *shared* social environment.

The relationship of *g* to sensory discrimination

In *Human Faculty*, Galton (1983) theorized that sensory acuity, especially the fineness of sensory discrimination, is a primal factor in mental development. He wrote '. . . the more perceptive the senses are of differences, the larger is the field upon which our judgment and intelligence can act', (p. 19). This was the essential premise on which he devised his laboratory tests of mental ability. But early on, Galton's simple sensory tests and the theory behind them were summarily dismissed by psychologists. Innumerable psychology textbooks have since proclaimed it 'common knowledge' that Galton was wrong and that general ability could be measured only by means of complex tests of reasoning, problem solving, practical judgment and general knowledge.

In recent years, though, there has been a revival of Galtonian research on the nature of mental ability. It shows that the same *g* factor found in complex mental tests is also found in the comparatively simple types of tests proposed by Galton. Possibly the general factor extracted from a large battery of such Galtonian tests would afford as good a measure of *g* as our standard IQ tests, provided the tests of sensory discrimination were of sufficient number and variety in every sensory modality to maximize their *g* and minimize their specificity. For example, in one study (Li, Jordanova & Lindenberger, 1998) two tests of tactile 'discrimination' (for texture and curvature) and one test of tactile sensitivity or 'acuity' (for pressure), when factor-analysed among a diverse battery of fourteen complex cognitive tests, showed a composite *g* loading of +0.46. In another study (Acton & Schroeder, 2001) correlations between psychometric *g* and scores on a test of colour discrimination ranged from +0.30 to +0.41. Measures of auditory pitch discrimination show similar correlations with *g* (Lynn, Wilson & Galt, 1989). In each of these sensory modalities, fineness of discrimination is consistently more *g*-loaded than sheer acuity, or the strength of a stimulus at the threshold of conscious detection. (The article by Li *et al*. (1998) provides a superb entrée to the contemporary literature in this vein.)

Although Galton's sensory tests now bear out his prediction that they correlate with general mental ability, his supposition that sensory discrimination is the essential basis of individual differences in mental development is less consistent with the totality of the present evidence, which Li *et al*. (1998) argue is better explained by the 'common-cause' hypothesis. This holds that simple sensory processes and the much more complex cognitive functions required by IQ tests are both manifestations of a third factor, namely, the speed and efficiency of neural information processing in the brain.

Mental processing speed

This theory brings us to another of Galton's creative ideas: that the *speed* of information processing is the basis of general ability. This idea motivated Galton's interest in the measurement of individual differences in reaction time (RT) to simple visual and auditory stimuli. He obtained RT data on over 9000 persons. Each one paid three pence to be tested in Galton's Anthropometric Laboratory in London's South Kensington Museum during the 1880s. Subsequent research has shown that

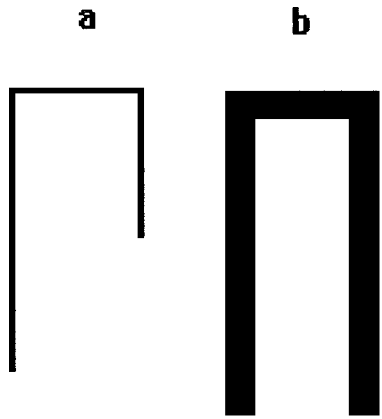

Fig. 2. The figure labelled **a** is presented for a brief duration (e.g. 70 msec) and is immediately followed by figure **b**, which covers and 'masks' the retinal after-image of figure **a**. The time interval between the onset of **a** and the onset of **b** needed for a person to report with 97·5% accuracy whether the longer 'leg' of figure **a** was on the left or the right side is termed that person's 'inspection time', or IT.

Galton's intuition about the relationship of information processing speed to general ability was scientifically a good lead.

Inspection time (IT)

The amount of time it takes to register a simple discrimination also correlates with *g*. The paradigm for testing this, called 'inspection time' (IT), yields measures in both the visual and auditory modes (Nettelbeck, 1987). In visual IT, the subject is repeatedly presented a stimulus, as shown in Fig. 2, for extremely brief, variable durations measured in milliseconds, followed by a 'masking' stimulus. On each trial the longer 'leg' appears randomly on either the left or the right side. The subject can take any amount of time to indicate on which side the longer 'leg' had appeared. What is measured is not the subject's response time, but the duration of the stimulus needed to make a correct response. Among young adults, making the correct decision on 90% of the trials requires, on average, a stimulus duration of about 70 msec.

The most recent review of the published evidence for a correlation between IT and conventional measures of mental ability, such as IQ, is provided by Mackintosh (1998, pp. 247–250), who summarized the evidence as follows:

> 'There can be no doubt that there is a significant correlation between IT and IQ. The precise value of that correlation may be a matter of some dispute . . . An upper-bound estimate of the unadjusted correlation would be no more than 0·50, and that with only some kinds of IQ test. But since the lower-bound estimate (at least for some kinds of IQ test) would have to be not much less than 0·25, we are left with a moderate, significant correlation, which is probably substantially higher that the correlation between RT (reaction time) and IQ,' (p. 248).

Although IT correlates more highly with those tests that also reflect, besides *g*, a non-verbal group factor, reflected by tests of visual–perceptual ability (Deary &

Crawford, 1998), evidence that IT is not exclusively a visual–perceptual ability is shown by the fact that *auditory* IT is also correlated with psychometric intelligence (Raz *et al.*, 1993). It is measured by two tones, easily distinguishable as either of low or high pitch, presented consecutively with each trial immediately followed by a masking tone of an intermediate pitch. The order of the high and low tones is random, and the person must report which tone, the first or the second, was the higher pitched. Subjects of lower IQ require longer durations of the auditory stimuli to make this discrimination.

Reaction time (RT)

Quickness of information processing would seem to be an advantage in early hominid survival and hence could be regarded as a Darwinian 'fitness character'. Galton therefore included it in his battery of tests as another elemental indicator of general ability. But Galton's own findings with RT looked unimpressive. Because of technical inadequacies in his method of measuring RT and the lack of any statistical method for determining the significance of his findings, his theory was discredited and abandoned during the first three-quarters of the 20th century. Recent research, however, has shown that Galton's idea was right after all. We now know for sure that measures of reaction time in what are called elementary cognitive tasks (ECTs) do in fact reflect the *g* factor (Vernon, 1987; Jensen, 1998a).

As an example, only one of the many ECTs that have been used to measure RT in this research will be described. The author has used it extensively in his research because of its elegant simplicity and its theoretically provocative results (Jensen, 1987).

Figure 3 shows the apparatus used to measure 'reaction time' (RT) and 'movement time' (MT). The subject presses the centrally located 'home' button, then a one-second preparatory stimulus sounds ('beep'), and after a random interval of one to four seconds, the reaction stimulus (a green under-lighted push-button six inches away from the home button) goes 'on'. The subject's task is to touch the green button as quickly as possible to turn off the light. Either one, two, four or eight lights can be presented, any one of which will go 'on' at random. The degree of uncertainty or 'information load' of the task is given by the binary logarithm of the number of response alternatives, so that 1, 2, 4 and 8 buttons represent 0, 1, 2 and 3 'bits' of information. (In information theory, a 'bit' is the unit of information that reduces uncertainty by one-half.) Two intervals in the subject's performance are measured in milliseconds: (1) RT, i.e. the time interval between the onset of the green light and the subject releasing the home button; and (2) MT, i.e. the time interval between the subject releasing the home button and touching the lighted green push-button, which instantly turns off the light. Figure 4 shows the results of this test when twenty trials are given on each of the four conditions of uncertainty (i.e. 0 to 3 bits of information). The graph has the same features for an individual (with response times averaged over twenty test trials) that we see in this graph based on the average of many people. The main features to be noted in these data are:

(1) Both RT and MT are very short time intervals. In fact, they are faster than the speed of conscious awareness of an external stimulus, which averages about 500 msec. Thus the RT is much too short to permit what we would normally think

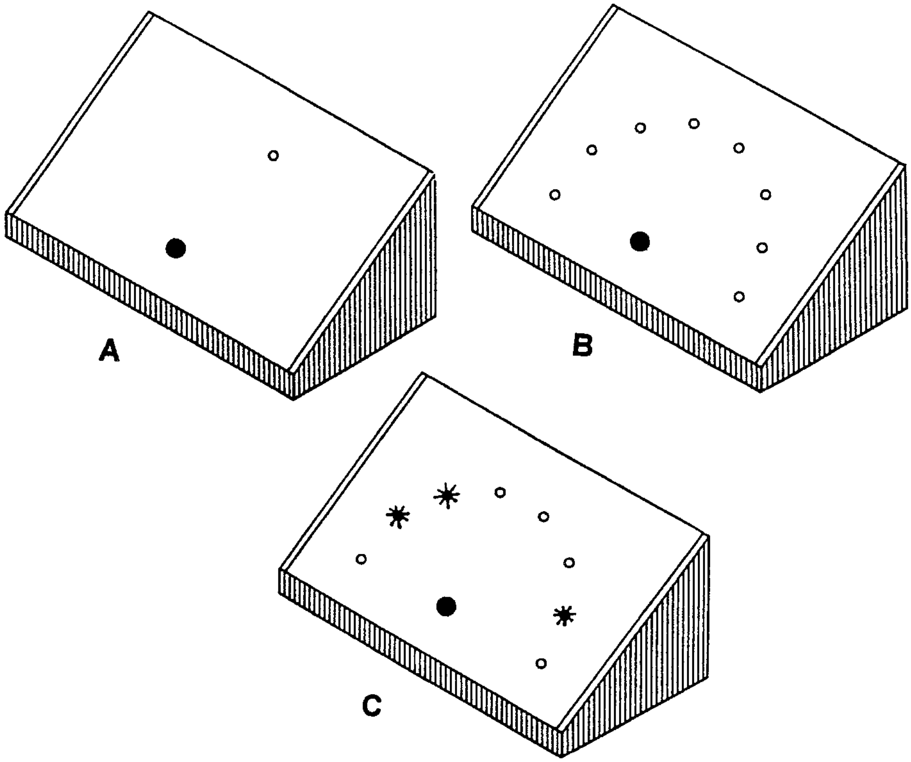

Fig. 3. The subject's response consoles for the Hick and the odd-man-out RT paradigms: (A) simple RT, (B) choice RT, (C) discrimination RT (odd-man-out). The black dot is the 'home' button; the open circles are green under-lighted translucent push-buttons six inches from the home button. In A and B only one button lights up on each trial; in C, three buttons light up simultaneously on each trial, with unequal distances between them; the subject is instructed to touch the 'odd-man-out', i.e. the one button that is farthest from the two other buttons.

of as reasoning or cogitation. Also, the task involves nothing one could regard as intellectual content or knowledge, and it calls for only minimal sensory–motor skill.

(2) To everyone's surprise, RT is much slower than MT. In testing some 2000 subjects the only exceptions to this finding were severely retarded persons with IQs below 50.

(3) Reaction time increases in a perfectly linear way as a function of the 'information load' (or conversely, uncertainty) in the stimulus array, a phenomenon known as Hick's Law. Movement time, however, does not follow Hick's Law; it is independent of the information load. When these data are factor-analysed, RT and MT show up on separate uncorrelated factors. Reaction time reflects information-sensitive cognitive processes; movement time reflects psychomotor ability.

The relationship of RT and MT to *g* is measured by 'untimed' or 'non-speeded' complex IQ tests such as Raven's Matrices and the Wechsler scales. Testing procedures have assured that individual differences in speediness of test-taking *per se*

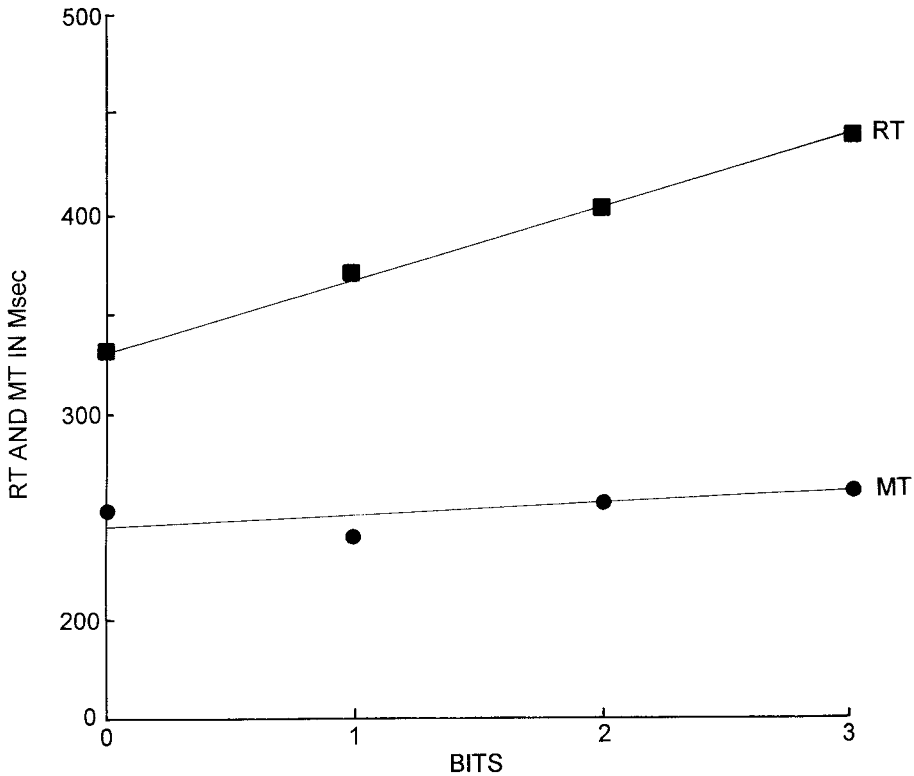

Fig. 4. Reaction time (RT) and movement time (MT) as a function of the 'information load' (scaled in 'bits' = $\log_2$ number of response alternatives) in the stimulus array, an effect known as Hick's Law (Jensen, 1987).

are not the source of the correlation between RT and IQ or *g*. The correlation between IQ and RT is always negative, of course, because persons with higher IQ generally have shorter RT than those with lower IQ. It is difficult to summarize briefly the magnitude of the RT–IQ correlation, however, because it varies considerably depending on many conditions that differ across various studies, such as the complexity or cognitive demands of the particular RT task and the range of IQ in the subject sample (which often consists of university undergraduates, who are mostly in the top quartile of the IQ distribution in the general population). Reaction time tasks have been variously elaborated to vary their cognitive load over a wide range, by increasing their demands on discrimination, judgment, attention span and retrieval of information from short-term or long-term memory (Caryl *et al.*, 1999). The more complex the task, up to a point, the longer the RT and the higher its correlation with IQ. The now considerable research evidence permits a number of fairly secure empirical generalizations.

(a) The negative correlation between RT and IQ is found in widely differing groups of subjects selected from all levels of the IQ distribution, going from the mildly retarded (IQs 60 to 80) to university students (IQs above 110), and members of Mensa, the high-IQ society (IQs above 132). These groups span nearly the whole range of the normal distribution of IQ in the population.

(b) Movement time has no independent correlation with IQ. A hierarchical factor analysis perfectly separates RT and MT as separate factors in a college sample (Carroll, 1991). The occasional slight positive correlation of MT with RT results mainly from those few subjects who leave the home button before ‘programming’ their movement, so their decision and movement times are confounded to some extent. (An apparatus has been devised to control this ‘leakage’ of some part of the decision time [i.e. RT] into the MT.) Also, the RT response itself has a very small motor component (i.e. lifting one’s index finger from the home button).

(c) As previously mentioned, the RT–IQ correlation increases when the information-processing demands of the RT task are increased. Even very small increments in the cognitive load slightly increase the RT’s correlation with IQ. For example, young adults process, on average, about 30 bits of information per second, and in the Hick task the increasing RT–IQ correlations for 0, 1, 2, 3 bits are −0·18, −0·19, −0·22, −0·23. As these correlations are based mostly on university students with a restricted range of IQ, they most likely underestimate the correlations that would be expected in the general population.

The correlation of median RT with IQ can be further increased if the information-processing demands of the task are increased. The ‘odd-man-out’ test (shown in Fig. 3C) requires a discrimination of the one light that is more distant from the other two lights when all three are lit up simultaneously. This increased stimulus-processing demand increases the subject’s RT by about 200 msec over the subject’s RT when only one light out of the eight alternatives is lit up, and it almost doubles the correlation of RT with IQ.

Many other elementary cognitive tasks (ECTs) have been used to measure processing time for retrieving information from immediate or short-term memory or long-term memory, and RT in all ECTs is correlated with untimed complex *g*-loaded psychometric tests involving reasoning, problem solving and acquired knowledge and cognitive skills. When the RTs on various ECTs are all combined, the composite measure correlates almost as highly with IQ tests as different IQ tests correlate with each other. Moreover, the *g* component of RT is the ‘active ingredient’ in these correlations. The degree of correlation between RT and other cognitive tests is directly related to the size of the respective tests’ *g* loadings. If we statistically remove *g* from a test, it no longer correlates with RT. So Galton’s idea that speed of information processing is an important aspect of individual differences in general intelligence has proved correct.

(d) Two other variables derived from the Hick task are also correlated with IQ independently of subjects’ median RT, viz. the trial-to-trial variability of RT and the slope of median RTs as a function of bits of information. When combined with median RT, the three variables together have a multiple correlation with IQ of about 0·35 or so, depending on the range of IQ in the subject sample.

Reaction time variability

When a person is given twenty or more trials on an RT task, we can reliably measure this person’s intra-individual variability as the standard deviation (SD) of the

RTs over a specified number (*n*) of trials, a variable labelled RTSD. It is usually more highly correlated (negatively) with *g* than is the subject's median RT for the same number of trials (Baumeister, 1998; Jensen, 1992). Higher IQ subjects are characterized by a smaller RTSD, that is, a higher trial-to-trial consistency of their RTs. It has been suggested that RTSD reflects the amount of 'neural noise' in the brain's information processing functions. It can be likened to the noisy static on a telephone line that slows communication when the conversing parties have to repeat words and phrases to get their messages across. Thus the inverse of RTSD reflects what can be called 'neural efficiency'.

Slope of reaction time as a function of information-processing load

The slope of RT as a function of bits of information is negatively correlated with IQ. For each successively greater increment in information load, lower IQ subjects have proportionally longer RTs than higher IQ subjects. Consequently, as the information load increases, the correlation between RT and IQ increases approximately linearly, a relationship that is typically underestimated because of failure to correct for certain statistical artifacts that are unique to the correlation between median RT and the slope of RT over bits (Jensen, 1998b).

(e) From early childhood to early adulthood RT becomes increasingly faster. It gradually slows down throughout the latter half of one's life span. Brain size follows a similar trajectory, as do psychometric measures of *g*. Movement time is less affected by age than is RT. Contrary to popular belief, age takes a greater toll on the cognitive, or *g*, demands of task performance than on its sensory–motor demands. But there are marked individual differences in the rate of decline in *g* with increasing age, those starting with the most generally having the advantage.

Racial differences in *g*

Based on his observations in Africa and scant other evidence, Galton conjectured that the difference in general ability between African Blacks and European Whites is twice the semi-interquartile range (equivalent to 1·33 standard deviations) between the means of the normalized frequency distributions of the two populations. In view of the inadequate evidence and the lack of actual psychometric measurement at that time, this pure conjecture was hardly convincing.

Galton's leading successor in differential psychology, the father of classical test theory, and the discoverer of *g*, Charles Spearman (1927), further hypothesized that the variation in the magnitude of the White–Black difference on diverse mental tests is a positive function of the tests' *g* loadings (Jensen, 1985). Spearman's hypothesis has been tested on nineteen independent sets of data based on nearly 100 different psychometric tests and representative samples of over 290,000 subjects (Blacks 15% and Whites 85%) (Jensen, 1998a; Nyborg & Jensen, 1999). In every study, Spearman's hypothesis was borne out and no data have yet been found that contradict the hypothesis. Figure 5 shows one such study. It is especially telling because the Black and White subjects were drawn from the same school classrooms and matched pair-wise for age, sex, grade in school and their parents' socioeconomic status (Naglieri & Jensen, 1987).

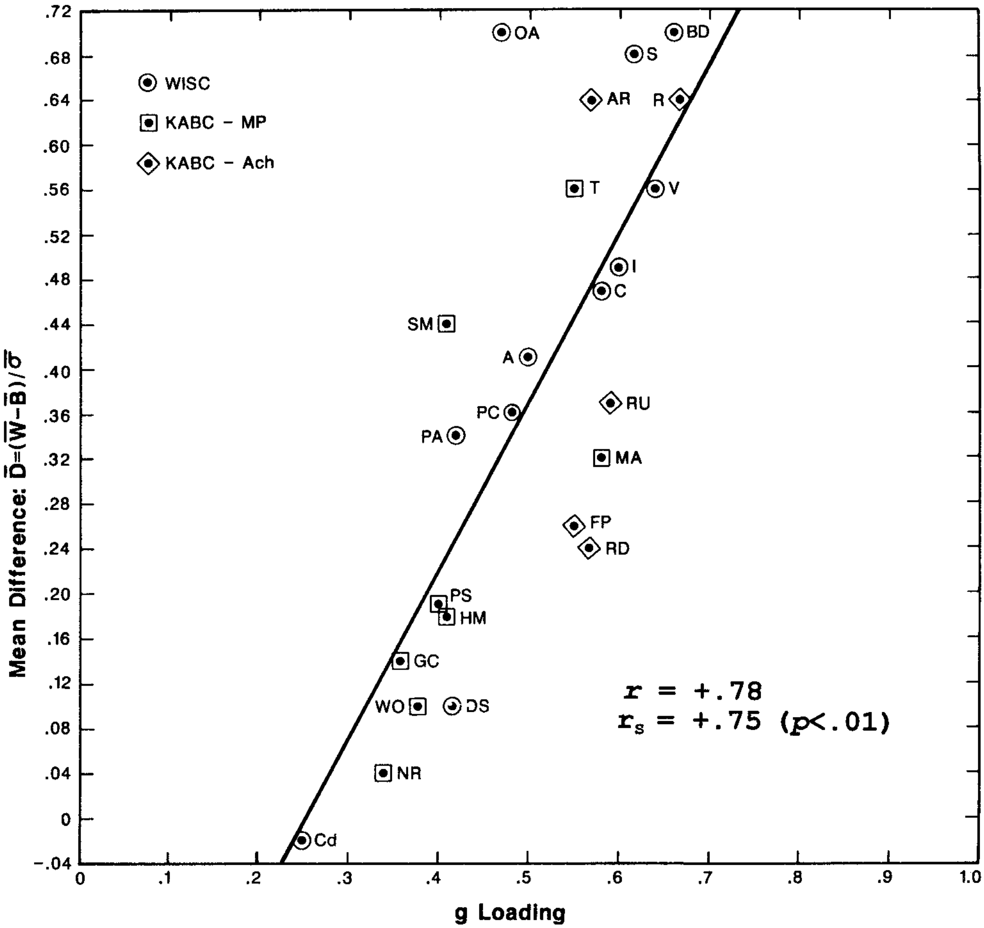

Fig. 5. Mean White–Black differences (scaled in the mean within-groups SD units) on all of the 25 sub-tests of the Wechsler Intelligence Scale for Children-Revised (WISC) and the Kaufman Assessment Battery for Children (KABC) (Naglieri & Jensen, 1987).

Galton's notion that the average White–Black difference in general ability, represented as Spearman's *g*, is 1·33 SD can now be evaluated with all of the data available in 1998, based on 149 psychometric tests given to samples totalling 43,892 Blacks and 243,009 Whites. The regression of the White–Black difference (in SD units) can be extrapolated on tests' *g* loadings to the point at which a hypothetical test would have the maximum possible *g* loading of 1·00. On such a hypothetical pure measure of *g*, the White–Black difference would be 1·31 SD, or equivalent to 20 IQ points (Jensen, 1998a, pp. 376–379). This is the best estimate we have of the White–Black difference in psychometric *g*. One may wonder if the proximity of Galton's conjecture of 1·33 SD to the present finding of 1·31 SD represents an amazing perspicacity or an extraordinarily lucky surmise.

Physical correlates of *g*

The belief that normal variation in general ability involves physical attributes of the brain is implicit in Galton's belief that such individual differences are largely hereditary. His technical means for empirically testing the relationship between brain and behaviour was limited to the measurement of head size as a crude estimate of brain size. He measured the average head size of Cambridge University students and found a positive relationship between head size and overall scholastic grade average (Galton, 1888), thus being the first to investigate the hypothesis that head size, as a proxy for brain size, is related to general mental ability. Many subsequent studies have confirmed Galton's conclusion that brain size and general ability are related (Rushton & Ankney, 1996). A number of recent studies in which brain size is precisely measured *in vivo* by means of magnetic resonance imaging (MRI) show correlations of +0.30 to +0.40 between brain size and IQ. Correlations in this range, of course, imply that other properties of the brain besides just its size are involved, given that the heritability of adult IQ is about 0·70. Although the correlation of IQ with brain size is now firmly established, what is still unknown is the causal basis of this correlation. Is it the number of neurones, or of synaptic connections, or the amount of myelin on the neural axons, or the density of glial cells (which constitute 90% of the total brain mass), or neurotransmitters, or some combination of such elements?

We now have the necessary technology with which to demonstrate a definite relationship between certain brain variables and psychometric *g*. But how these variables actually cause *g*, or even whether they do, is unknown at present. We are still in the inductive phase of this science, and in the search for causes it is essential to establish a network of correlations as dependable guideposts in our search (Jensen, 1997b; Jensen & Sinha, 1993).

The 'method of correlated vectors' is one technique that can tell us whether any given variable is related to the *g* factor independently of non-*g* sources of variance in a given battery of mental tests (Jensen, 1998a). This procedure is illustrated in Fig. 6 by means of a bivariate scatter diagram, which is one of Galton's several statistical inventions. Shown in this example is the correlation ($r = +0{\cdot}80$) between the *g* loadings of the eleven sub-tests of the Wechsler Adult Intelligence Scale (WAIS) and each of these sub-tests' correlations with a measure derived from the amplitude of the evoked potential (EP), called the 'habituation' of the EP (Schafer, 1985).

The EP measures the change in electrical potential in the brain's response to an external stimulus. The evoking stimulus in Schafer's study was the sound of a sharp 'click.' The subject sits passively and hears a 'click' every two seconds. The brain's EP is recorded from an electrode attached to the vertex of the scalp. The EP gradually habituates; that is, the amplitude of the evoked 'brain wave' decreases over the series of 50 successive 'clicks'. The habituation index is the mean difference in amplitude of the evoked potential between the first and second sets of 25 'clicks'. It was correlated +0.59 with IQ. Figure 6 shows that the correlation between habituation of the EP and IQ is due to its *g* component. Generally, the larger the IQ sub-test's *g* factor loading, the higher is its correlation with the EP habituation index.

Other factors besides *g* might add or subtract from the correlation between a given brain variable and IQ or might have no effect at all. In Schafer's study, three other

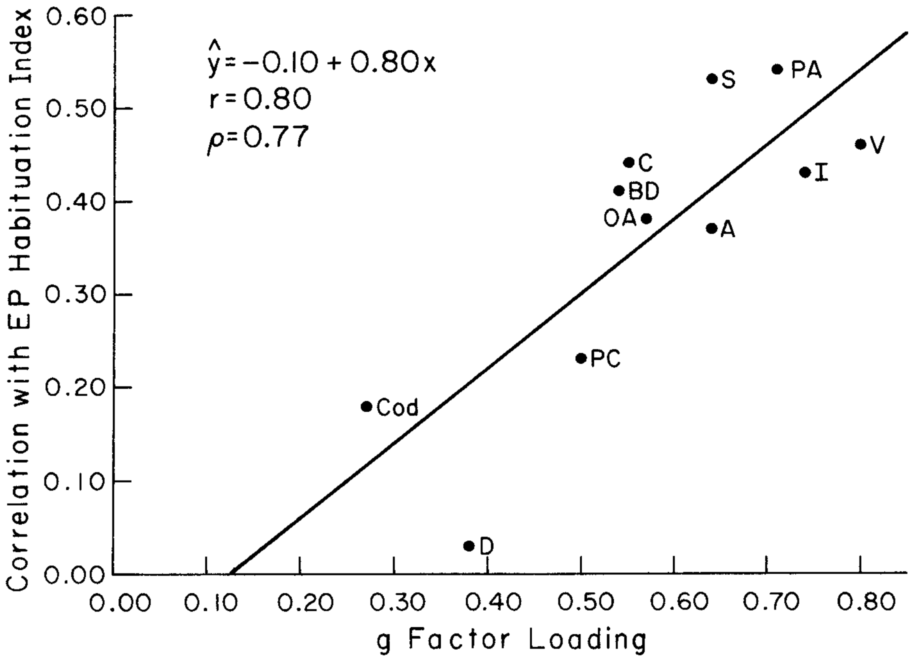

Fig. 6. Scatter diagram showing the correlation (*r*) between the *g* loadings of the various sub-tests of the WAIS battery and these sub-tests' correlations with the habituation of the evoked potential (EP).

factors independent of *g* (verbal, spatial and memory), that were extracted from the WAIS battery, had near-zero correlations with the EP. When the *g* component was statistically removed from the each of the WAIS sub-tests, they had near-zero correlations with the EP habituation index.

The method of correlated vectors has also shown that the degree of correlation between each of various mental tests and each of the following variables is significantly predicted by the degree to which each of the tests is loaded on the *g* factor independently of other sources of variance in the given psychometric battery, with the typical vector correlations in parentheses (references to the individual studies given in Jensen, 1998a):

- The *heritability* of the tests (0·70).
- *Assortative mating*: the correlation between spouses' test scores (0·95).
- *Inbreeding depression* of test scores (0·80).
- *Heterosis*: outbreeding elevation of test scores in the offspring of inter-racial mating (0·50).
- *Reaction time* on various ECTs (0·80).
- *Intra-individual variability* in RT on various ECTs (0·75).
- *Head size* as a correlated proxy for brain size (0·65).
- *Brain evoked potentials*: habituation of amplitude (0·80).
- *Brain evoked potentials*: complexity of waveform (0·95).
- *Brain intracellular pH* level (0·63).

- *Cortical glucose metabolic rate* during psychometric testing, measured by PET scan (0·79).

Nothing could be more important for the advancement of this research than to secure replications in different laboratories of the findings listed above. This is especially needed for those studies based on measurements of evoked electrical brain potentials, usually referred to as Event-Related Potentials (ERPs). Some of the techniques for demonstrating a correlation between ERPs and IQ or other psychometric measures of *g* have not yet been brought under sufficient experimental control or standardized procedures to dependably yield consistent results across independent studies. It seems most unlikely that the numerous, though often inconsistent, results reported in the literature are all purely chance effects. The puzzling inconsistencies, however, leave this body of research hardly interpretable in theoretically coherent terms. This is especially true of measures of the complexity of the waveform of the averaged ERP. (A more detailed discussion of the problems with this measure in relation to IQ is provided by Mackintosh, 1998, pp. 236–242.) Yet one can hardly ignore the remarkable relationship ($r = +0{\cdot}95$) found in the one and only study that looked at the correlation between the vector of the twelve Wechsler sub-test *g* loadings in one study and the vector of correlations of the ERP complexity measure with each of the sub-tests. Could this be merely a chance phenomenon? It is unlikely. A major effort to replicate this result consistently is required. Dependable replications in different laboratories, if successful, would indeed be a major advance.

Replications are also required of a number of other studies performed in laboratories around the world that have shown significant correlations between single highly *g*-loaded tests and various brain variables such as MRI-measured brain size, EEG coherence, event-related desynchronization of brain waves, and frontal lobe alpha brain wave frequency.

The geneticist T. E. Reed was the first to hypothesize that the genetic basis of *g* results in part from individual differences in brain nerve conduction velocity (NCV; Reed, 1988). This hypothesis led to collaboration in a study which demonstrated a relationship between non-verbal IQ scores (on Raven's matrices) and NCV in a brain tract from the retina to the visual cortex (Reed & Jensen, 1992). The results, shown in Fig. 7, have been essentially replicated in a recent study by Andres-Pueyo, Bonastre & Rodriguez-Fornells (1999) at the University of Barcelona. Only future replications of the brain correlates of *g* can provide a firm basis for the further exploration and development of *g* theory. The heritability of *g* absolutely implies some physical substrate in the brain. Discovering its nature remains a major frontier of research in differential psychology.

The future of the Galton paradigm

Galton's ideas have long been a rich fount of continuing research in differential psychology. This paper has described his ideas about human mental ability and their most recent developments. Overall, Galton's paradigm, with its roots in evolutionary theory, genetics and physiology, has proved essentially sound. Its fruit now represents mainstream scientific understanding of individual and group differences in mental ability. Most of the major but as yet unanswered questions that have arisen in the

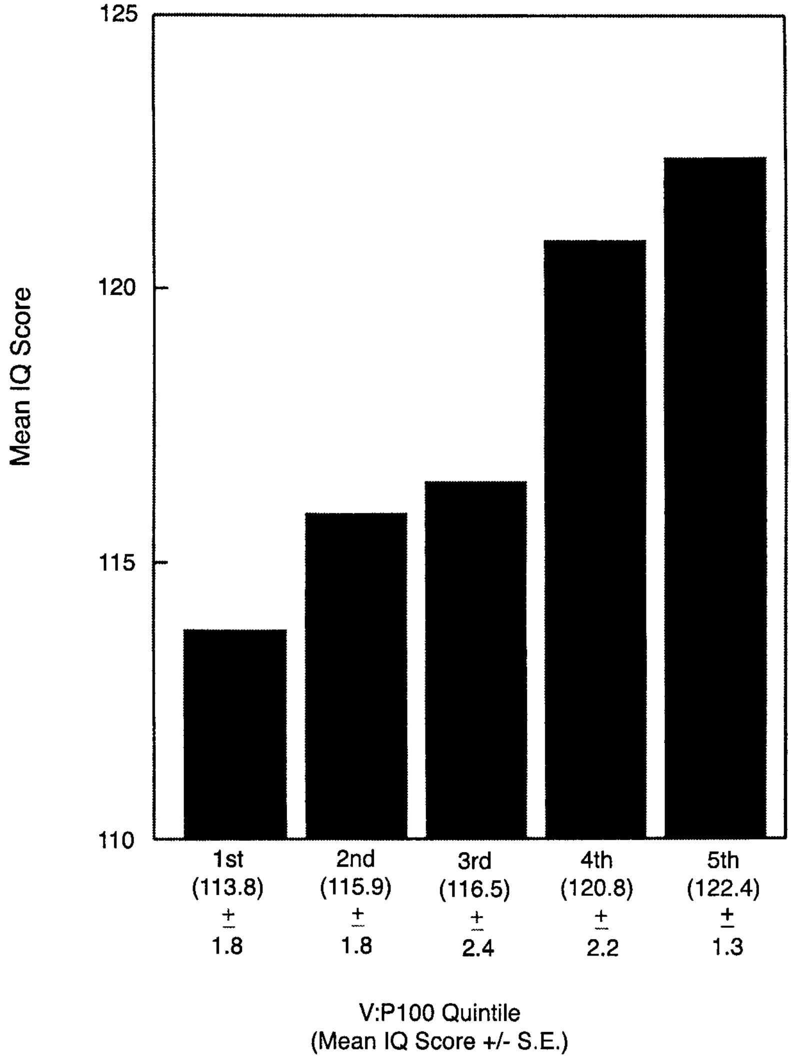

Fig. 7. IQ means (in parentheses) ± standard errors of each quintile of the distribution of nerve conduction velocity (NCV) measured in the visual tract in 147 male college students. This restricted IQ sample comprises only the top one-third of the IQ distribution in the general population. Individual values of velocity (V:P100) were based on the P100 latency of the visual evoked potential. The measures of NCV in this sample range from the slowest at 1·75 m/sec to the fastest at 2·22 m/sec.

modern renaissance of the Galtonian paradigm are essential grist for further explorations in molecular genetics, brain science and evolutionary psychology. These fields are presently pursuing their unique contributions to furthering our understanding of human mental ability.

A fundamental question that now confronts investigators arises from an important conceptual and empirical distinction between the two main facets of the effort to understand general ability.

- First, there are the specialized design features of the brain, its neural structure, circuitry and functions that make information processing possible. Information processing comprises the whole class of behaviour we refer to as 'intelligence': attention, perception, discrimination, generalization, learning, memory, language, thinking, reasoning, relation eduction, inference, problem solving and the like. The neural components of these various functions have different loci in the brain and some depend on specialized modules. Although these brain modules originated through the evolutionary process in the prehistoric past, some of their unique functions can be adopted for serving entirely novel modern demands that played no part in natural selection. However, it is highly unlikely that the evolutionary process has resulted in a super-module that serves as a general problem solver. Rather, the author suggests that the explanation of *g* will probably have to be sought in other than modular structures. It is more likely to depend on chemical or neural mechanisms that affect all the various modules that play a role in specific intellectual functions, causing them all to be positively correlated.
- Second, there are individual differences in the speed and efficiency of operation in all these various functions and their diverse manifest abilities. These individual differences are correlated across many different abilities. These positive correlations essentially are the basis for the concept of *general* ability or *g*.

The key question, then, is what are the properties of the brain that cause all these diverse types of information processing to be positively correlated? In brief, what is the cause of *g*? This is one of the big questions in science today. Answering it will engage many researchers for quite some time. Is the causal basis of *g* something functionally different and separate from the neural design features or circuitry of the brain that make human intelligence possible in the first place?

The few brain correlates of *g* discovered so far are quantitative differences that affect the speed and efficiency of information processing, such as brain size (number of neurones?, or myelin? or glia cells?), neural conduction velocity, and brain glucose metabolic rate. Individual differences in such quantitative aspects could affect many of the brain's modular functions in common. From an evolutionary viewpoint, it seems a reasonable hypothesis that, among members of one and the same species, ubiquitous correlations between various brain modules and their associated abilities more likely result from heritable physiological and biochemical conditions that affect many of the brain's functions in common, rather than from individual differences in the brain's design features and neural circuitry (Jensen, 1997b).

The questions generated by the Galtonian paradigm are clearly posed. In view of recent developments in genetics, brain science, and evolutionary psychology, it seems predictable that the causes of *g* will be discovered sooner rather than later in the new millennium. It is much less predictable how society then will regard this knowledge or will put it to use. If the zeitgeist is imbued with the likes of Sir Francis Galton's vision, prudence and humane spirit, the future can be entrusted hopefully to the unrelenting application of humanity's evolved intelligence.

Acknowledgments

The author especially wishes to thank Dr Leslie Jones of The Galton Institute for his generous discussions while preparing this 1999 Galton Lecture and which have materially enhanced its contents.

References

Acton, G. S. & Schroeder, D. H. (2001) Sensory discrimination as related to general intelligence. *Intelligence* **29**, 263–271.

Andres-Pueyo, A., Bonastre, R. M. & Rodriguez-Fornells, A. (1999) Brain nerve conduction velocity, magnetic nuclear resonance and intelligence: New data. Paper presented at the *9th Biennial Convention of the International Society for the Study of Individual Differences*, Vancouver, BC, July 1999.

Baumeister, A. A. (1998) Intelligence and the 'personal equation'. *Intelligence* **26**, 255–265.

Bouchard, T. J., Jr (1996) Behaviour genetic studies of intelligence, yesterday and today: The long journey from plausibility to proof (The Galton Lecture). *J. biosoc. Sci*. **28**, 527–555.

Burt, C. (1962) Francis Galton and his contributions to psychology. *Brit. J. stat. Psychol*. **15**(1), 1.

Carroll, J. B. (1991) No demonstration that *g* is not unitary, but there's more to the story: Comment on Kranzler & Jensen. *Intelligence* **15**, 423–436.

Carroll, J. B. (1993) *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. Cambridge University Press, Cambridge.

Caryl, P. G., Deary, I. J., Jensen, A. R., Neubauer, A. C. & Vickers, D. (1999) Information processing approaches to intelligence: Progress and prospects. In: *Personality Psychology in Europe*, Vol. 7, pp. 181–219. Edited by I. Mervielde, I. Deary, F. de Fruyt & F. Ostendorf. Tilburg University Press.

Chorney, M. J., Chorney, K., Seese, N., Owen, M. J., Daniels, J., McGuffin, P., Thompson, L. A., Detterman, D. K., Benbow, C., Lubinski, D., Eley, T. & Plomin, R. (1998) A quantitative trait locus associated with cognitive ability in children. *Psychol. Sci*. **9**, 1–9.

Colom, R., Garcia, L. F., Juan-Espinoza, M. & Abad, F. J. (in press) Null sex difference in general intelligence: Evidence from WAIS-III: Implications for psychological assessment. *Spanish J. psychol.*

Colom, R., Juan-Espinoza, M., Abad, F. J. & Garcia, L. F. (2000) Negligible sex differences in general intelligence. *Intelligence* **28**, 57–67.

Deary, I. J. & Crawford, J. R. (1998) A triarchic theory of Jensenism: Persistent, conservative reductionism. *Intelligence* **26**, 273–282.

Galton, F. (1869/1892/1962) *Hereditary Genius: An Inquiry into its Laws and Consequences*. Macmillan/Fontana, London.

Galton, F. (1883/1907/1973) *Inquiries into Human Faculty and its Development*. AMS Press, New York.

Galton, F. (1888) Head growth in students at the University of Cambridge. *Nature* **38**, 14–15.

Gottfredson, L. S. (Ed.) (1997) Intelligence and social policy [special issue]. *Intelligence* **24**(1).

Jensen, A. R. (1980) *Bias in Mental Testing*. Free Press, New York.

Jensen, A. R. (1983) Effects of inbreeding on mental-ability factors. *Person. indiv. Diff*. **4**, 71–87.

Jensen, A. R. (1985) The nature of the black–white difference on various psychometric tests: Spearman's hypothesis. *Behav. Brain Sci*. **8**, 193–219.

Jensen, A. R. (1987) Individual differences in the Hick paradigm. In: *Speed of Information Processing and Intelligence*. Edited by P. A. Vernon. Ablex, Norwood, NJ.

Jensen, A. R. (1992) The importance of intraindividual variation in reaction time. *Person. indiv. Diff.* **13**, 869–881.

Jensen, A. R. (1997a) The puzzle of nongenetic variance. In: *Heredity, Intelligence, and Environment*. Edited by R. J. Sternberg & E. L. Grigorenko. Cambridge University Press, Cambridge.

Jensen, A. R. (1997b) The neurophysiology of *g*. In: *Processes in Individual Differences*. Edited by C. Cooper & V. Varma. Routledge, London.

Jensen, A. R. (1998a) *The g Factor*. Praeger, Westport, CT.

Jensen, A. R. (1998b) The suppressed relationship between IQ and the reaction time slope parameter of the Hick function. *Intelligence* **26**, 43–52.

Jensen, A. R. (2000) Charles Spearman: The discoverer of *g*. In: *Portraits of Pioneers in Psychology*, Vol. IV. Edited by G. A. Kimble & M. Wertheimer. American Psychological Association, Washington, DC.

Jensen, A. R. & Sinha, S. N. (1993) Physical correlates of human intelligence. In: *Biological Approaches to the Study of Human Intelligence*. Edited by P. A. Vernon. Ablex, Norwood, NJ.

Johnson, R. C., McClearn, G. E., Yuen, S., Nagoshi, C. T., Ahearn, F. M. & Cole, R. E. (1985) Galton's data a century later. *Am. Psychol.* **40**, 875–982.

Li, S-C., Jordanova, M. & Lindenberger, U. (1998) From good senses to good sense: A link between tactile information processing and intelligence. *Intelligence* **26**, 99–122.

Lynn, R. (1991) Race differences in intelligence. *Mankind Q.* **31**, 254–296.

Lynn, R. (1999) Sex differences in intelligence and brain size: A developmental theory. *Intelligence* **27**, 1–12.

Lynn, R., Wilson, R. G. & Galt, A. (1989) Simple musical tests as measures of Spearman's *g*. *Person. indiv. Diff.* **10**, 25–28.

Mackintosh, N. J. (1998) *IQ and Human Intelligence*. Oxford University Press, Oxford.

Naglieri, J. A. & Jensen, A. R. (1987) Comparison of black–white differences on the WISH-R and K-ABC: Spearman's hypothesis. *Intelligence* **11**, 21–43.

Nagoshi, C. T. & Johnson, R. C. (1986) The ubiquity of *g*. *Person. indiv. Diff.* **7**, 201–208.

Nettelbeck, T. (1987) Inspection time and intelligence. In: *Speed of Information Processing and Intelligence*. Edited by P. A. Vernon. Ablex, Norwood, NJ.

Nyborg, H. & Jensen, A. R. (1999) Black–white differences on various psychometric tests: Spearman's hypothesis tested on American armed services veterans. *Person. indiv. Diff.* **28,** 593–599.

Plomin, R. (1997) Identifying genes for cognitive abilities and disabilities. In: *Heredity, Intelligence, and Environment*. Edited by R. J. Sternberg & E. L. Grigorenko. Cambridge University Press, Cambridge.

Raz, N., Willerman, L., Ingmundson, P. & Hanlomnf, M. (1993) Aptitude-related differences in auditory recognition masking. *Intelligence* **7**, 71–90.

Reed, T. E. (1988) A neurophysiological basis for the heritability of intelligence. In: *Intelligence and Evolutionary Biology*. Edited by H. J. Jerison & I. Jerison. Springer, Berlin.

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.

Rowe, D. C. (1994) *The Limits of Family Influence: Genes, Experience, and Behaviour*. Guilford Press, New York.

Rushton, J. P. & Ankney, C. D. (1996) Brain size and cognitive ability: Correlations with age, sex, social class, and race. *Psychonomic Bull. Rev.* **3**, 21–36.

Schafer, E. W. P. (1985) Neural adaptability: A biological determinant of *g* factor intelligence. *Behav. Brain Sci.* **8**, 240–241.

SPEARMAN, C. E. (1904) General ‘intelligence’, objectively determined and measured. *Am. J. Psychol.* **15**, 201–293.

SPEARMAN, C. E. (1927) *The Abilities of Man: Their Nature and Measurement*. Macmillan, New York.

STIGLER, S. M. (1986) *The History of Statistics: The Measurement of Uncertainty Before 1900*. Belknap Press of Harvard University Press, Cambridge, MA.

VERNON, P. A. (ED.) (1987) *Speed of Information Processing and Intelligence*. Ablex, Norwood, NJ.