

Chapter 10

# *Phlogiston, Animal Magnetism, and Intelligence*

*Arthur R. Jensen*

*University of California, Berkeley*

Q: What do the terms in the above title have in common?

A: From a scientific standpoint, they all have proved unsatisfactory.

In 18th-century chemistry, *phlogiston* was a hypothetical element, the pure essence of fire that remained latent in combustible material and escaped as visible flames in the process of burning. Hence, materials were reduced in substance and lost weight as they burned, usually. Certain kinds of matter gained rather than lost weight in burning, an observation that complicated phlogiston theory, and made it necessary to hypothesize that phlogiston is a substance having a negative weight. Without any operational means of defining and measuring phlogiston, it was impossible for the early chemists to get an empirical handle on it for scientific study that could resolve disputes about its nature. Phlogiston theorists could not move beyond merely pointing at the various manifestations of fire; that is, the observable phenomena their theory was supposed to explain. Perennial arguments over the phlogiston concept, however, fueled the drive of early chemists to seek a better explanation for the obvious-

ly real and observable phenomenon of fire. When finally the nature of combustion was correctly explained in terms of rapid oxidation, the phlogiston theory was completely abandoned and is now just a quaint relic in the history of chemistry.

*Animal magnetism*, a theory put forth by Anton Mesmer, supposedly explained certain phenomena of what, in the 18th century, was called “mesmerism,” and later become known as “hypnotism.” By analogy with the “lines of force” that a magnet exerts at a distance on objects made of iron, Mesmer explained that a similar force, which he called “animal magnetism,” flowed from the mind of the hypnotist to that of his subject, inducing a trancelike state and allowing the hypnotist to take control of the subject’s subconscious mind and behavior. Numerous investigators, including Benjamin Franklin, could find no evidence for the existence of animal magnetism or any correlates of it beyond the particular hypnotic phenomena it tried to explain. In fact, so lacking was animal magnetism in any real explanatory power, it soon became no more than merely a synonym for hypnosis. The notion of animal magnetism was later supplanted by more fruitful theories that related hypnosis to other psychological phenomena and were couched in empirically testable terms.

The term *intelligence* shares many of the same scientifically unsatisfactory characteristics of phlogiston and animal magnetism. The parallels are so close that, as our science advances, we might expect intelligence to be abandoned as a concept in scientific discourse, as were phlogiston and animal magnetism. Although the word itself may eventually slip from the technical vocabulary of differential psychology and psychometrics, it will most likely survive in the popular vernacular. But I would prefer that, as psychologists, we immediately drop this phlogistonlike term in any scientific context.

How much more evidence do we still need that psychologists are unable to reach agreement on the meaning of intelligence after nearly a century of trying, unsuccessfully, to “define” it and “theorize” about it (see Jensen, 1987a; Sternberg & Detterman, 1986)? The operational definition originally suggested by Boring (i.e., “Intelligence is what intelligence tests measure”) in no way solves the problem, and is now generally recognized as patently vacuous, not to say fatuous. Shouldn’t this be sufficient warning that a “theory of intelligence” has about as much chance of success, scientifically, as phlogiston or animal magnetism? The hopelessly muddled concept of intelligence is at best useless and at worst a hindrance to efforts by behavioral and brain scientists who would advance investigation of what is obviously an important realm of phenomena, certainly one of the most important in the province of psychology. Abandoning the fruitless quest for intelligence in no way negates the actual phenomena of interest, any more than scraping phlogiston negated the phenomena of combustion. But making concessions to the recognized inadequacy of intelligence by mere lexical modifications surely does not go far enough. Confusion is only compounded by adopting the plural form of the word—intel-

ligences—or by attaching various adjectives (e.g., *verbal* intelligence, *practical* intelligence, *social* intelligence, *global* intelligence, etc.). Nor should we adopt any *new* term to substitute for intelligence. The word and the concept alike should go completely. As I hope to explain, scientific psychology can get along very well without intelligence.

And so, having discarded intelligence, including all its synonyms as well as the concept itself, what are we left with and where do we begin? Obviously, with the observable phenomena of interest, as must every science.

### THE EMPIRICAL PHENOMENON OF VARIANCE IN MENTAL ABILITIES

*Ability* can be defined objectively as anything a person can do with some consistency. But its definition must have three limiting aspects: (a) It is consciously or voluntarily initiated behavior, which excludes involuntary reflexes and behavioral effects resulting from emotional states, dreaming, trauma, fatigue, disease, or drugs; (b) It has some quantifiable reliability or temporal stability, that is, a better than chance repeatability (within a specified time frame) under similar circumstances; and (c) The behavior can be assessed in terms of an objective standard. *Objective* here simply means high agreement among observers or recording instruments. *Standard* implies some index of the performance in terms of accuracy (i.e., degree of approximation to a clearly defined criterion) or time taken (e.g., response latency). Jumping over a 2-foot hurdle, lifting a 20-pound weight, answering “35” to  $5 \times 7 = ?$ , spelling the word for “a large pachyderm that has tusks and a trunk,” pressing a button as quickly as possible (measured in milliseconds) when a particular light goes on, throwing a baseball  $x$  number of feet, solving a given problem in calculus, and parking one’s car parallel to the curb in a space between two other cars—each of these performances qualifies as an ability, assuming it can be done with better than chance consistency. Obviously, the number of abilities is virtually infinite, and one could make up “items” to “test” every known or conceivable ability. The number of possible tests would be limited only by the imagination and inventiveness of test makers. Thus the domain of abilities, by this definition, is completely open-ended, although bounded by the three criteria listed above.

A *mental* ability can also be defined objectively. First, it must meet the three limiting conditions for defining ability (see above). Second, it must be an ability for which an insignificant part of the total variance in the particular ability in a given population is associated with individual differences in sensory acuity or motor strength and dexterity per se. At least in principle, a mental ability can be demonstrated by some kind of performance that obviates or minimizes the role of any particular sensory or motor mechanism. A blind person,

for example, can understand a question that is spoken as well as a sighted person can, although the sighted person may also be able to read the question when it is presented in a written form. Understanding the question, as indicated by making an appropriate response to it, is what qualifies the demonstrated ability as *mental*; the modality of presentation and the particular effector mechanisms involved in responding are nonessential features of the demonstrated mental ability. Individual differences in a particular mental ability are demonstrable without their having any correlation with individual differences in sensory-motor functions per se.

### FROM ABILITIES TO TESTS AND FACTORS

It seems obvious that many abilities, tested by means that would permit their objective demonstration by an individual, would, when tested in a heterogeneous population, show individual differences, which may be expressed statistically as *variance*. If, for example, 60% of the adult population are able to jump over a 2-foot hurdle fairly consistently (say, on any 4 out of 5 trials) and 40 percent are unable to do so, this item of ability would have a variance of  $.6 \times .4 = .24$ . The population could similarly be divided by many other possible hurdle heights, each with some variance. And, in all likelihood, there would be some positive *covariance* (and hence *correlation*) among all the different hurdles, so the total variance for hurdle jumping in the population would be equal to the sum of the variances of each of the hurdles *plus* twice the sum of all their covariances. If there are many items (say,  $n$  items), and if they are positively correlated, twice the sum of all the item covariances, of which there are  $n(n - 1)/2$ , will be much greater than the sum of the item variances.

Now, this sum of all the covariance terms represents a low-level abstraction, which we could call variance in *general* hurdle-jumping ability. An individual's raw "score" would be, say, the height of the highest hurdle that could be cleared in 4 out of 5 trials, when, say, 100 hurdles are graduated in half-inch steps. Or the person's raw score could be the average height of each of the hurdles that was cleared on one trial. In either case, it should be noted that the variable of *general* hurdle-jumping ability in this case is an abstraction one step removed from the person's *ability* to jump a particular hurdle. It is no longer an *observed* ability (as here defined), but an inferred ability *factor*, albeit a quite low-order factor. By low-order I mean that it may have only quite limited generality in the whole domain of physical abilities. I should emphasize that there is nothing at all "good" or "bad" or in any other way *valuative* implied by the definitions of ability and of lower order and higher order factors—terms that imply only different degrees of *generality*. Generality, as applied to factors, simply refers to the number of abilities that are correlated with the factor.

A *factor* represents some degree of *association* between abilities, quantized by a *correlation coefficient*. The overall score (i.e., number of passing performances on the separate items) on a set of abilities in which all of them are highly correlated with one another forms a low-order factor. By obtaining individual measurements of a great many abilities in a population, we can form numerous sets, or clusters, each composed of the abilities that are the most highly correlated with one another. A single such collection of highly intercorrelated ability items constitutes a *homogeneous test*. Obviously, the degree of homogeneity will vary depending on the size of the correlation coefficients specified as the criterion for inclusion of a given ability item in the set. If the criterion correlation is relatively high, there will be many sets with relatively few ability items in each one, and each set will be quite homogeneous. If the criterion correlation is low, there will be fewer sets with more ability items in each one, and each set will be less homogeneous.

A set of positively intercorrelated ability items is called a *test*. An individual's *raw score* on the test is the number of items (or trials) on which the individual's performance passes some objective criterion or is expressed as the average (over items or trials) of some metrical variable, such as response time or speed. A test composed of ability items that meet the stated criteria of *ability* and *mental* is a *mental ability test*. Because a person's score on such a test is not the person's observed behavior on any particular ability item but is derived from a collection of such items, it represents an ability factor, which may be narrow or broad, depending on the test's homogeneity, so long as it has *some* degree of homogeneity.

The total variance of raw scores (or any linear transformation of them, such as standard scores) on a test so defined can be decomposed into *true score* variance and *error* variance. The proportion of the total variance that is true score variance is expressed as the test's *internal consistency reliability*, usually measured by the Kuder-Richardson formula (KR-20). It is monotonically related to the test's homogeneity or average interitem correlation, and, because the internal consistency reliability represents that proportion of the total variance composed of all the item covariances, it is an increasing function of the number of items in the test. The well-known Spearman-Brown formula is a precise expression of this relation between the number of items in a test and its internal consistency reliability. Hence a test can be made to be just as reliable, in this sense, as the test maker wishes it to be. Since the reliability coefficient asymptotically approaches unity by increasing the length of the test, practical considerations must dictate the optimum reliability for any given purpose.

The main purpose of this discourse so far is to remove any mystery that may be attached to the essential idea of a mental ability test and to show that such tests can be formed in a systematic fashion by assorting together correlated ability items from the theoretically unlimited pool of performances (poten-

tial test items) that meet the stated objective criteria for classification as “ability” and as “mental.”

It comes as a surprise to many people, including psychologists, that abilities, as here defined, are as *unique* as they in fact are. That is, ability items (hereafter referred to simply as *items*), if selected at random from a large and diverse pool, show quite low correlations with one another; correlations between such items are typically in the  $+0.05$  to  $+0.15$  range, averaging about  $+0.10$ . In the most internally consistent tests, those expressly constructed by psychometricians to maximize item homogeneity, the average item intercorrelation seldom exceeds  $+0.20$ . Yet, if there is a fairly large number of items in such a homogeneous test, its internal consistency reliability (i.e., the proportion of true score variance) can go well above  $.90$ . (A 100-item test with an average item intercorrelation of  $+0.20$  would have an internal consistency reliability of  $.99$ .) But it is important to realize that an overall score on a multiitem test is one step removed from an observable ability, which itself exists only at the level of item performance. Anything we may say beyond that level of observable behavior is an abstraction, or inferential. Thus even a single test score is an inference. Until it can be properly described or interpreted within an explicit objective framework of analysis, it is no more than a test score. What such scores may *mean* is a question that can only be answered empirically and at a number of different levels of analysis.

*Factor analysis*, of one type or another, is an essential method for classifying mental abilities or the tests made up of them, much as Mendeleev’s periodic table of the elements has been essential to chemistry. (This analogy, however, should not be carried much further.) Factor analysis is essentially a method for partitioning the total variance (or individual differences) of a large number of abilities (or tests) into a much smaller number of different “sources” of variance called *factors*, which may be correlated (*oblique* factor axes) or uncorrelated (*orthogonal* axes) with one another.

Without going into the methodology of factor analysis, suffice it to say that all of the correlations among a number of homogeneous tests can be analyzed into a smaller number of factors that represent sources of variance the tests have in common to varying degrees. Some of these sources of variance, or factors, are more general than others. The *generality* of a factor refers to the number of different tests in which it is represented and also to the amount of the total true score variance it accounts for in the whole collection of tests. Because the factors extracted from mental ability tests differ in generality, they can be thought of hierarchically, going from the least general factors at the bottom of the hierarchy to the most general at the top. Such a model is said to represent the factorial structure of abilities. It is important to emphasize that we are speaking here only of *covariance structures*, that is, the pattern of intercorrelations among individual differences in the measured abilities. At this point,

there is no reference to causal mechanisms. Covariance structures and factors per se have no direct implications concerning the nature of either the cognitive processes or the brain processes that mediate performance and are the ultimate loci of individual differences. A hierarchical factor structure, however, does suggest that individual differences in the cognitive processes and brain processes that underlie performance on mental tests most probably also differ in generality. The covariance structure of abilities, represented in the case of mental abilities by a hierarchical factor analysis, is simply a point of departure for empirical investigation of the cognitive processes and neural mechanisms that are responsible for the factor structure. Before arriving at that point, however, a few more points about the factor analysis of mental abilities are in order.

First of all, a hierarchical factor structure, indicating different levels of generality, could not adequately represent the covariance structures in the mental abilities domain were it not for an important empirical fact, namely, the phenomenon of *positive correlation among all mental abilities*. This phenomenon is also called *positive manifold* when it is represented in terms of a correlation matrix of all positive correlations. I have found no evidence of any two or more mental abilities that are consistently uncorrelated or negatively correlated in a large unrestricted or random sample of the population. The few observed exceptions to this most important empirical generalization are entirely explainable in terms of measurement error, sampling error, biased sampling of the population, restriction of the range of ability in the sample, and inclusion of test items which represent types of performance that do not meet the essential criteria for a mental ability. The phenomenon of positive manifold in mental abilities is one of the most important facts to be explained by any theory of human mental ability.

Second, we can divide up the total variance obtained in a large battery of tests in a number of steps, as follows:

1. At the item level: (a) the sum of the *item variances*, which constitutes the *error variance*, and (b) the sum of the *item covariances*, which constitutes the *true score variance*.

2. The total true score variance can be divided into (a) *common factor variance*, or the variance that different items or tests have in common, and (b) *specificity*, or that proportion of a test's true score variance that is not common (i.e., uncorrelated) with any other tests in the battery.

3. The total common factor variance can be hierarchically analyzed into a number of factors of varying degrees of generality, from narrow to broad, with the number of factors decreasing at each level, thus forming a triangular structure. At the lowest level are (a) *first-order* factors (also called *primary* factors or *group* factors). These first-order factors are closely identified with different types of tests, such as verbal, numerical, spatial, and memory, to name a few. The fact that even the most homogeneous of such tests are all positively corre-

lated with one another indicates that there are more general, or higher order, factors that they have in common. All factors beyond the first-order factors are termed *higher order* factors. After the first-order factors, at the next higher level of generality are (b) *second-order* factors, which represent very broad categories of tests, such as those that depend largely on previously learned knowledge or skills (i.e., so-called *crystallized* ability) as contrasted with those that involve the solution of relatively novel problems (i.e., so-called *fluid* ability). These second-order factors are also correlated, indicating a third-order level of generality. Although there is usually only one third-order factor, namely, the single most general factor of the matrix, it is possible, with enough tests of sufficient diversity, to obtain two or three third-order factors. Even higher order factors beyond the third are a theoretical possibility, but they seldom occur in factor analyses of mental ability tests. Usually, at the third-order (occasionally at the fourth order), only a single factor emerges, called the *general factor*, or *g* (also known as “Spearman’s *g*” and “psychometric *g*”). This *g* factor, which is the apex of the factor hierarchy, has the greatest generality of any factor, in the sense that it is represented in every test. Also, it often accounts for more of the total common factor variance than any of the lower order factors, or, in some cases, even more than all of them combined.

It would take us too far afield to describe the various mathematical models and methods by which a hierarchical factor analysis can be performed. The main methods in current practice, however, yield quite comparable solutions. The two main types of analysis can be described briefly as (a) “top-down,” in which the *g* factor is extracted first and the remaining common factor variance is analyzed into a number of primary factors; and (b) “bottom-up,” in which the correlations between the first-order factors are themselves factor analyzed to yield second-order factors, and the process is repeated until *g* emerges at the highest level. It is rare that both methods do not yield the same factors.

It is mathematically possible to scatter and submerge the variance of the *g* factor in the several primary factors and to constrain these factors to be perfectly orthogonal (i.e., uncorrelated), so that no higher order factors can be extracted. This is accomplished by an orthogonal rotation of the original factor axes, using, for example, the Varimax criterion. In such a case, the *absence* of higher order factors (including *g*) is a purely mathematical artifact. A strong argument can be made that orthogonal rotation of the first-order factors, such as Varimax, is entirely inappropriate when applied to the factor analysis of abilities. In fact, any method that would hide the *g* factor (i.e., by distributing its variance among all of the primary factors and mathematically forcing them to be uncorrelated) when a general factor actually exists in the correlation matrix, is simply wrong. To elaborate this argument properly would involve technical issues that are beyond the scope of this chapter (see Jensen & Weng,



in press). Here I can only emphasize that a bottom-up method of hierarchical factor analysis cannot possibly yield a  $g$  factor that is not actually latent in the zero-order correlations among the original variables (or tests, in this case). Therefore, it is the best method in an exploratory factor analysis of abilities.

To summarize, the results of factor analysis make it possible to represent the total variance of a given test as comprising the following components:  $g$  + one or more first-order (and possibly high-order) factors + specificity + error. (The first two terms together are called the test's *communality*; the last two terms together are called the test's *uniqueness*.) Factor analysis yields quantitative estimates (i.e., proportions of total variance) of these terms for all of the tests subjected to the factor analysis. Like any other statistics, the estimates are generalizable, with a determinable standard error and with reference to the population on which they are based.

It is important to consider two entirely independent sources of sampling error in factor analysis: (a) *subject sampling* from some specified population, as in every statistic; and (b) *psychometric sampling*, or the selection of mental tests that enter into a factor analysis. The latter is more problematic, because there is no extant "population" of all possible mental tests from which we can draw random samples. The theoretical population of mental tests is unlimited, and investigators are free to invent whatever varieties of mental tests they please, so long as the items meet the stated criteria for mental ability. When we draw more or less random samples of tests from the entire catalog of existing tests, however, a hierarchical factor analysis always yields a  $g$  factor, and the  $g$  is remarkably similar from one sample of tests to another. The number and nature of other factors besides  $g$  that emerge depends on the number of types of abilities or homogeneous tests (e.g., verbal, numerical, spatial, etc.) that enter into the factor analysis. The  $g$  factor, on the other hand, always emerges provided there are a fair number of tests and enough variety among the tests to allow the extraction of other factors besides  $g$  to prevent the  $g$  from being heavily admixed with any particular group factor. Poor psychometric sampling, in this respect, results in a somewhat "contaminated"  $g$  factor; the "impurities" can be removed by adding more tests to the battery that will form additional primary factors.

In innumerable factor analyses of mental ability tests, the ubiquity of  $g$  and the remarkable constancy of  $g$  across so many different batteries of diverse tests strongly suggest that theoretically there exists a "true"  $g$ , whereas the  $g$  extracted from any given battery of tests is simply a statistical estimate of it, in the sense that the "true" measurement of any quantitative variable can be conceived of as the mean of an unlimited number of measurements, each of which is an estimate of the true value. (By "estimate" is meant the true value plus random error.) Hence the  $g$  factor of abilities especially commands scientific interest.

## THE RELATIVE IMPORTANCE OF THE *G* FACTOR

Psychometric *g* is both more mysterious and more challenging for scientific study than most other factors, because it cannot be described in terms of any particular knowledge content, skill, type of test, or observable behavior. It is whatever causes even the most dissimilar tests, to all outward appearances, to be positively correlated. It is intuitively rather easy to see why two verbal tests are correlated, or two tests of spatial visualization, or two of numerical manipulation. It is harder to understand why dissimilar tests, say, vocabulary and block design, are correlated; or such dissimilar tests as choice reaction time, backward digit span, and pitch discrimination. The amazing fact that all of these tests (and innumerable others) are positively correlated is reflected in their *g* loadings. The general factor extracted from a variety of measures of conscious or voluntary *learning* (which may exclude certain forms of classical conditioning of autonomic and involuntary reflexes) is factor analytically indistinguishable from the *g* of psychometric tests (for a recent review of this literature, see Jensen, 1988).

A major task of any theory of mental ability is to explain *g*, which is the most important and central fact about human abilities. But other factors are also essential features of a comprehensive map of human abilities. Their number is unknown and theoretically indeterminable, but a number of distinct group factors, independent of *g*, are now well established by a multitude of studies. (This literature has been comprehensively reviewed by Carroll, 1993.) Every highly replicable factor eventually must be studied in its own right and explained in a theoretical context, much as *I* and others have been trying to do in the case of *g*. It is important to study the nature of these non-*g* factors *independently* of *g*, that is, with *g* held constant statistically or by subject selection. The explanations of various factors and their properties may differ considerably, of course.

Before considering a theoretical explanation of *g*, I should mention some of the most salient facts about its relation to other variables, most of which I have discussed in detail in other publications listed in the present bibliography.

1. In the *predictive validity* of tests—for scholastic achievement, college grades, job performance, occupational category, success in various armed forces training schools—the chief “active ingredient” is *g*. Removing *g* from tests (or their validity coefficients) that have demonstrated their practical utility would render them virtually useless, because it is the *g* factor that accounts for most of the practical predictive validity of the tests used for educational and employment selection (Jensen, 1993a; Ree & Earles, 1992).

2. A necessary corollary of the first point is that schooling, academic performance, job performance, and various occupational categories are themselves *g* loaded to varying degrees. The size of their *g* loading depends on the extent to which they involve types of performance that qualify as mental abilities, as

previously defined. Hence,  $g$  is more predictive of success in training and performance in most professional and managerial occupations than in types of work that make less complex demands on mental abilities, such as manual labor and unskilled jobs. Creating selection tests that minimize  $g$  would be the surest way to damage their validity for any real-life criteria involving mental abilities (Schmidt, Ones, & Hunter, 1992).

In types of work that involve special talents and particular highly developed skills, such as musical, literary, and artistic performance,  $g$  usually acts as a *threshold* variable. That is, the probability of successful development of the special talent falls off precipitously for individuals who fall below some critical or threshold value on  $g$ . Hence it would be exceedingly unlikely to find the full range of  $g$  in any random sample of, say, musically accomplished performers or composers. In fact, just about every kind of occupation has a critical threshold on the distribution of  $g$ , although this threshold differs markedly for different occupations. The IQ is a rough index of  $g$ , and occupational categories differ markedly in the lowest IQs found among persons who are employed in the various occupations, indicating differing thresholds on the scale of  $g$  for successful performance in different occupations (Jensen, 1980a, Chap. 8).

3. Psychometric  $g$  is also related to a number of variables completely outside the province of mental tests or any variables that are thought of as abilities. One way of showing this is to obtain the correlations,  $r$ , between a number of different mental tests and some other nontest variable (call it  $x$ ), and then to determine the rank-order correlation,  $\rho$ , between (a) the set of correlations of the various tests with  $x$  and (b) the  $g$  loadings of the various tests. (Both the correlations and  $g$  loadings are corrected for attenuation, so variation in the tests' reliability will not systematically affect the rank-order correlation between a and b.) For various batteries of tests, such as the 12 subtests of the Wechsler Intelligence Scale for Adults, very substantial correlations, ranging from about +.50 to +.95, have been found between variables a and b (above) when variable  $x$  is one of the following: (a) the heritability coefficient (i.e., the proportion of genetic variance in test scores), (b) the correlation between a number of different genetic kinships on the various tests, (c) the correlation between spouses on the various tests, (d) the magnitude of inbreeding depression of test scores, and its genetic counterpart, (e) outbreeding enhancement (heterosis) of test scores, seen in the offspring of racial crosses, (f) habituation of the amplitude of auditory evoked cortical potentials recorded from a scalp electrode affixed at the vertex of the skull, (g) a measure of the complexity of the wave form of the average evoked potential, and (h) various paradigms of choice and discrimination reaction time measurements.

One of the probably important ways that  $g$  differs from all lower order factors is this: These  $x$  variables (listed above) that are related to  $g$  are found to be unrelated, or in some cases only slightly related, to other well-established

ability factors *independent* of *g*, such as verbal and spatial factors. (More detailed descriptions and references to these various studies are in Jensen, 1987b.)

Total scores on standard IQ tests, which are quite highly *g* loaded, are also correlated with a host of physical variables, such as height, weight, brain size, certain blood antigens, serum uric acid level, vital capacity, basal metabolic rate, myopia, asthma and various allergies, and a number of other physical variables. (This literature has been reviewed by Jensen & Sinha, 1993.) Also the rate of glucose metabolism in certain regions of the brain, as measured by positron emission tomography (PET scan), is correlated (negatively) with scores on Raven's Progressive Matrices, a highly *g*-loaded nonverbal test involving inductive and deductive reasoning (Haier et al., 1988).

The origins of these correlations between *g* and various physical variables are only scarcely understood, if at all. But the evidence leaves no doubt that the population variance on mental ability tests reflects latent variables, predominantly *g*, that are profoundly enmeshed with organismic structures in complex ways. A comprehensive theory of abilities must eventually account for the observed relationship between *g* and these anatomical and physiological variables. Some of these connections between *g* and physical characteristics have undoubtedly come about in the course of evolution, whereas others may reflect environmental effects, such as nutrition, that affect certain physical variables, including the neural anlage of abilities, during ontogenetic development. Also, certain of these correlations between *g* and physical traits are due in large part to cross-assortative mating; that is, persons of above-average *g* selecting mates who are above-average in, say, height, or physical attractiveness, or any other visible features popularly deemed desirable in any culture that also values the salient achievements associated with a higher level of *g*. A methodology, based on statistical manipulations of sibling data, that assists in analyzing the nature of all these kinds of physical-mental correlations, has been explicated elsewhere (Jensen, 1980b; Jensen & Sinha, 1993). Study of the relationship between *g* and physical or other nonpsychometric variables seems quite germane to research on the biological evolution of mental abilities. For example, the relation between inbreeding depression and *g*, according to genetic theory, suggests that *g*, more than any other mental ability factor, reflects a fitness character in the Darwinian sense; that is, it has been subject to natural selection during some period in the course of human evolution (Jensen, 1983).

### **EMPIRICAL FINDINGS GERMANE TO THE EXPLANATION OF *G***

Attempts to explain *g* in terms of the information content or specific skills involved in ability measures or tests must be completely dismissed. These fea-

tures are merely the vehicles for the ordinal measurement of individual differences in  $g$  (Jensen, 1992a). We increasingly approximate the “true”  $g$  (analogous to a “true” score in classical test theory) as we add more and more measures of diverse mental abilities to the correlation matrix. Hence, as the obtained  $g$  factor more closely approximates the hypothetical true  $g$ , it is increasingly stripped of those properties that can be described in terms of the specific characteristics of the tests. What is reflected by  $g$ , ultimately, is individual differences in some general property or quality (these should be stated in the plural as well) of the brain. The brain, of course, is the one and only organ that, if made nonfunctional, would preclude any kind of behavior that meets the definition of a mental ability. A theory of mental ability, therefore, must ultimately be a theory of the brain, its anatomical structures, and neurophysiological processes. Psychometrics and experimental cognitive psychology, however, provide important hypotheses and techniques for research at the interface of brain and behavior.

### REACTION TIME (RT) IN ELEMENTARY COGNITIVE TASKS (ECTs)

When I began my search for the causal underpinnings of  $g$ , I harkened back to one of the earliest hypotheses, originally suggested more than a century ago by Sir Francis Galton, to the effect that individual differences in general ability are to a large degree due to differences in the speed of brain processes, which are reflected in the speed of simple mental activity such as reaction time (RT) to an external stimulus (called the reaction stimulus, or RS). My research over the past decade, as well as that of many other investigators, on the relation between RT and  $g$  has amply proven that Galton’s hypothesis is essentially correct, with certain qualifications.

RT is an especially useful technique in this type of research, because it permits highly reliable ratio–scale measurement of the speed with which a person can perform extremely simple mental tasks, called *elementary cognitive tasks* (ECTs), that are within the capability of virtually all persons over a wide age range who are not afflicted by gross sensory-motor or neurological impairments. RT also has the virtue of permitting comparison of performances in a considerable variety of ECTs measured on a common scale of real time (typically measured in milliseconds). ECTs are specially devised to tap one or more hypothesized *information processes* (also called *cognitive processes*) presumed to be necessary for performance of the particular ECT, such as stimulus apprehension, encoding of a stimulus and discrimination between stimuli, retrieval of information from short-term memory (STM) or from long-term memory (LTM), choice, decision, or other mental manipulation of the input, response selection, and response execution, to list some of the hypothesized information

processes. In my chronometric laboratory, we have studied many of these ECTs, especially those involving the fewest and simplest processes, such that the average RT is usually in the range of 200 to 600 msec and seldom as long as 1000 msec on any ECT. It turns out that the most interesting correlations with  $g$  are found with those ECTs to which most subjects respond in this range of relatively short RT (i.e., 200–600 msec). More complex tasks, resulting in longer RT, apparently leave more room for idiosyncratic variation in cognitive strategies or other vagaries of performance than the simpler ECTs. These erratic sources of variance in RT actually *reduce* the correlation between RT and psychometric  $g$ . The most likely reason for this is that  $g$  itself does not reflect individual differences in strategies or idiosyncratic aspects of problem solving, but reflects individual differences in the speed of certain elementary mental operations and their neural basis.

Without reviewing the many studies of the relation between psychometric  $g$  (or scores on one or another psychometric test that is highly  $g$ -loaded) and RT measured in many different ECTs, I shall briefly mention the principal findings that seem most important for a theory of  $g$ . (More comprehensive reviews and additional references are provided in Jensen, 1982, 1987b, 1987c, 1992b, 1993b.)

1. In simple RT (i.e., a single response to a single stimulus) very little information processing is involved, the only uncertainty in the task being the precise time when the reaction stimulus (RS—a green light going on) will occur at random within the 4-second interval following the preparatory stimulus (a beep). RT is the interval between the onset of the RS and the subject's removing his finger from the pushbutton (called the "home" button). Movement time (MT) is the interval between release of the home button and touching the RS (green light), which turns it off.

Simple RT generally shows correlation with  $g$  between zero and  $-.20$ , averaging about  $-.10$ . The correlation is usually smaller in high  $g$  groups and, controlling for reliability, is larger in young children and in the mentally retarded. Our explanation for this lies in the fact that simple RT comprises two main components: (a) a *peripheral* component consisting of sensory lag, or stimulus transduction, motor nerve conduction time, and muscle lag; and (b) a *central* component due to information processing in the brain, involving neural conduction time and synaptic delays. The peripheral component of simple RT, which is not directly involved in information processing, constitutes a relatively large proportion of the total variance in simple RT. As the peripheral component does not reflect speed of information processing, its variance attenuates the correlation between simple RT and  $g$ .

2. Choice and discrimination RTs, of course, involve the same peripheral component as simple RT, but the greater complexity of the choice and discrimination RT tasks requires more information processing, which is reflected in the longer RT in these more complex forms of RT and in the greater vari-

ance attributable to central processes. Hence RT to more complex ECTs has higher correlations with psychometric  $g$ . For single ECTs, these correlations are generally in the  $-.20$  to  $-.50$  range, averaging about  $-.35$ . By combining the RTs from several different ECTs, the correlations with  $g$  go up to about  $-.70$ ; that is, about half of the variance in psychometric  $g$  is accounted for by a composite of RTs on a variety of ECTs, any of which can be performed in less than one second by the vast majority of subjects (usually college undergraduates). Moreover, the information content of these ECTs consists of nothing that could be called “intellectual” in the ordinary sense of that term, and if the ECTs were taken as nonspeeded tests and the responses scored “right” or “wrong,” there would be zero variance in the populations studied. In fact, prior to administering certain ECTs, potential subjects are screened with an untimed paper-and-pencil version of the ECT and those who miss a single item are dismissed. Hence the RT variance on these ECTs reflects individual differences in the speed of information processing rather than in acquired information content.

3. When simple RT is removed from various forms of choice or discrimination RT, either by simple subtraction or by statistical partialling, the correlation between the difference (i.e., choice RT–simple RT) and psychometric  $g$  is larger than are the correlations for either simple or choice RT alone. In other words, the peripheral component that simple RT shares with choice RT acts as a suppressor variable in the correlation between choice RT and  $g$  (Jensen & Reed, 1990).

4. An experimental manipulation that increases the RT– $g$  correlation consists of presenting subjects with a dual task; that is, while the information of one task is being held in STM, the subject has to perform some RT task. For example, the subject is presented a series of five or six digits to memorize in 3 seconds, then the subject must perform a choice RT task, and finally, the subject must repeat the memorized digits. Under this condition, both tasks, digit span and choice RT, will each show a higher correlation with  $g$  than when either task is performed separately. The dual task paradigm, which has been studied most extensively in relation to  $g$  by Stankov (1988) and Vernon (1983), suggests that a concept of STM *capacity*, in which there are individual differences, must be a necessary ingredient of the explanation of  $g$ . If by straining the capacity of *working memory* (WM) the rank-order correlation between RT and  $g$  is significantly increased, it necessarily means that some additional source of variance besides sheer processing speed is involved—variance associated with the capacity of WM. A theory of individual differences in WM capacity, therefore, is a necessary adjunct to a theory of  $g$ . I will say more about it later on.

5. Another aspect of RT that must be taken into account is *intraindividual* variability in RT across trials, measured as the standard deviation ( $SD$ ) of RTs on  $n$  trials, henceforth called SDRT. It would simplify matters if SDRT were

completely redundant with the mean or median RT over trials, but it is not (Jensen, 1992c). Although RT and SDRT are highly related, with correlations ranging between about  $+.5$  and  $+.7$  in different samples and for different ECTs, it turns out that RT and SDRT are not perfectly correlated after correction for attenuation, and when both are entered into a multiple regression equation to predict psychometric  $g$ , each variable makes an independent contribution to the multiple correlation. SDRT usually makes the larger contribution, despite its considerably lower reliability. With correction for attenuation, SDRT almost always correlates more with  $g$  than the median RT on the same set of trials. In other words, high  $g$  persons, as well as having faster overall RTs, also have more consistent RTs (hence smaller SDRT) from trial to trial. In a study of simple RT, for example, the mean SDRT for 46 mildly retarded young adults was 108.1 msec, for 218 vocational college students 48.8 msec, and for 280 university students 29.8 msec (Jensen, 1982, Table 1). (The groups also differ in the coefficient of variation, that is, SDRT/RT: .23, .14, .10, respectively.)

For all subjects, the longer RTs in a given number of trials have the largest variance and are the most highly correlated with  $g$ , but the higher correlation is not simply an artifact due to their larger variance (Kranzler, 1992; Larson & Alderton, 1990). In brief, there is an intrinsic relation between individual differences in RT variability, or SDRT, and  $g$ , independently of the average RT. So there are these two elements, median RT and RTSD, that must enter into a theory of  $g$ . An individual's median RT could be said to reflect the *speed* of information processing, or of the neural transmission of information in the brain, while SDRT could be said to reflect *oscillation* or random "noise" in the transmission and processing of information.

6. A direct measure of nerve conduction velocity (NCV) in the central nervous system is theoretically valuable for establishing that speed per se is an element of  $g$ . It could be the case, for example, that the speed of processing reflected in RT is merely a derivative of the intraindividual trial-to-trial variability of RT. Since there is some physiological limit to the speed of reaction (somewhere around 170 msec), and if everyone has pretty much the same physiological limit, large individual differences in SDRT could arise only by the production of a certain number of relatively long RTs, more for some subjects than for others. The resulting skewness of individuals' distributions of RTs would, of course, produce corresponding differences in the means or medians of the RT distributions, and thus median RT would be merely a derivative of the intraindividual variability in RTs, and thus only intraindividual variability, rather than speed per se, would be responsible for the correlation of RT with  $g$ . That this is *not* the case is shown by the correlation between direct measurements of the speed of neural conduction, or NCV, in a single nerve tract and  $g$ .

Trying to find a physiological basis for the considerable genetic heritability of general mental ability, Reed (1988) hypothesized that nerve conduction



velocity (NCV) is the causal factor. To test this hypothesis, short latency visually evoked potentials (VEPs N70 and P100) in response to pattern-reversal stimulation and recorded over the primary visual cortex were obtained on 147 male college undergraduates. The latencies of the earliest clearly defined neural impulses transmitted from the retina through the visual tract to the visual cortex are quite short—only 70 to 100 msec. Dividing the individual's head length by the mean latency of his VEP gives an estimate of NCV. These approximate measures of NCV (labeled V:N70 and V:P100) were significantly correlated with IQ scores on Raven's Advanced Progressive Matrices, a highly *g*-loaded, nonspeeded, nonverbal test of complex reasoning ability. The correlation between Raven IQ and NCV was  $+0.18$  ( $p = .025$ ) for V:N70 and  $+0.26$  ( $p = .002$ ) for V:P100. Correction for restriction of range of IQ in the college sample raises these correlations to  $+0.27$  and  $+0.37$ , respectively, and correction for attenuation (which was not attempted) would raise the correlation to perhaps as high as  $+0.50$  (Reed & Jensen, 1992). Figure 1 shows the mean IQ within each quintile of the V:P100.

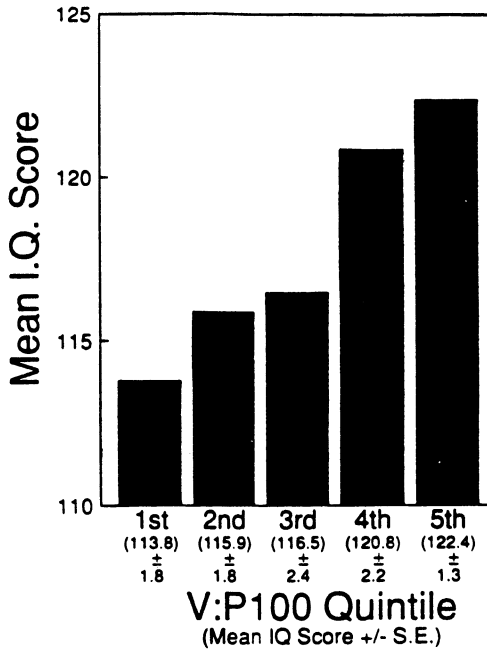
This finding means that the speed of neural transmission in a single, well-defined nerve tract that involves no more than four synapses and that is not a derivative of intraindividual variation in VEP latencies is correlated with a measure of *g* based on a nonspeeded, self-paced test of complex reasoning.

The theoretical significance of this finding extends to another issue as well; that is, the question of whether individual differences in such basic neural processes as NCV cause individual differences in the higher mental processes involved in *g* (i.e., the "bottom-up" hypothesis) or vice versa (the "top-down" hypothesis).

The "bottom-up" hypothesis holds that there are stable individual differences in relatively simple but pervasive neural processes, such as NCV and synaptic delay, which govern the speed and efficiency of information transmission in the whole central nervous system, and that these properties are involved at all levels of information processing, from relatively simple tasks, such as choice RT, to the much more complex problems in conventional IQ tests. Because individual differences in these neural properties are involved at all levels of information processing, individual differences in, say, choice RT are correlated with scores on complex psychometric tests.

The "top-down" hypothesis, on the other hand, holds that individual differences in higher level mental processes, strategies, and various other metaprocesses that are obviously involved in the kinds of problem solving seen in most highly *g*-loaded psychometric tests are also solely responsible for individual differences in RT in relatively simple ECTs, and that this top-down influence accounts for the correlation between performance on ECTs and the complex problem solving in psychometric tests.

The "top-down" hypothesis is contradicted by the finding of a correlation between NCV in the visual tract and *g*. The latency of the neural impulses in



**Figure 1.** Distribution of mean IQ scores in V:P100 quintiles. The distribution of V:P100 values (i.e., the NCV based on the P100 latency) of the 147 students, from the lowest NCV (1.75 m/sec.) to the highest (2.22 m/sec.) was divided into quintiles. Quintile 1 contains the 20% of students with the values, quintile 2 contains the 20% of students with V:P100 values between the 20th and The linear regression of individual IQ on quintile number (1,2 ...) has a slope of 2.21 IQ points per quintile, with no significant deviation from linear trend.

---

From "Conduction Velocity in a Brain Nerve Pathway of Normal Adults Correlates with Intelligence Level," by T. E. Reed & A. R. Jensen, 1992, *Intelligence*, 16, 259-278. Reprinted by permission.

the visual tract (recorded over the visual cortex) is much shorter than the total amount of time needed for neural impulses to reach the higher cortical centers involved in solving Raven Matrices problems, and in fact it is even much shorter than the time needed for a subject to gain conscious awareness of an external stimulus. Therefore, the VEP latencies cannot be controlled by the higher mental processes.

The explanation for the observed correlation between NCV in the visual tract and  $g$  is based on the reasonable hypothesis that, since the neurons in the visual tract and in the cortex share a common origin and have common features (e.g., small caliber axons and similar conduction speeds), they are very

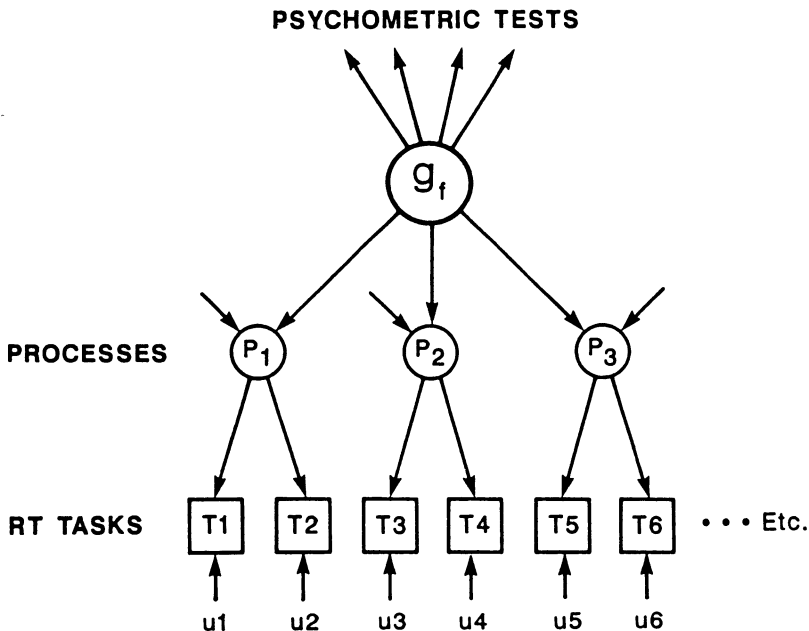
similar, and thus individual differences in visual tract NCVs and cortical NCVs are correlated. Because information is transferred from one cortical region to another via axons at some velocity and across synapses with some delay, the mean NCV and cumulative synaptic delay would affect the speed of information processing at every level of cognitive complexity. Individual differences in mean cortical NCV, therefore, appear to be a basic component of  $g$ .

Certain structural design features of neuronal organization or architectonics are probably involved in some of the major group factors independent of  $g$ —what Spearman referred to as the specialized “engines” of the brain in which there are distinct individual differences in addition to individual differences in the properties they all share in common, such as NCV, and that account for  $g$ . It is here taken for granted that specific neural structures with complex functional organization and patterning are essential for information processing at any level. But at present there is a dearth of empirical knowledge of just how or to what degree these design features of the cortex contribute to individual differences in cognitive abilities. However, we do have some evidence now that NCV in the brain may alone account for as much as perhaps 25% of the  $g$  variance in the general population (Reed & Jensen, 1992). To become a pillar in the theory of mental ability, of course, this finding will need ample replication. If it holds up, it would be a crucial step indeed toward understanding variation in human mental ability.

### **Why the Apparent Ceiling on the RT- $g$ Correlation?**

It is common knowledge in RT research that RT based on any particular ECT seldom correlates more than about .3 to .4 with  $g$ , and that the multiple correlation based on the RTs from a number of ECTs designed to measure different cognitive processes seldom exceeds about .6 or .7, when corrected for attenuation. This ceiling on the RT- $g$  correlation may lead us to think that some fairly large part of the  $g$  variance, perhaps as much as half of it, must be due to some source of individual differences besides the mental speed variable reflected in RT. This additional source of  $g$  variance has been attributed to differences in knowledge base, attentional resources, motivation, problem-solving strategies, executive processes, and other metaprocesses, to name the most frequently mentioned.

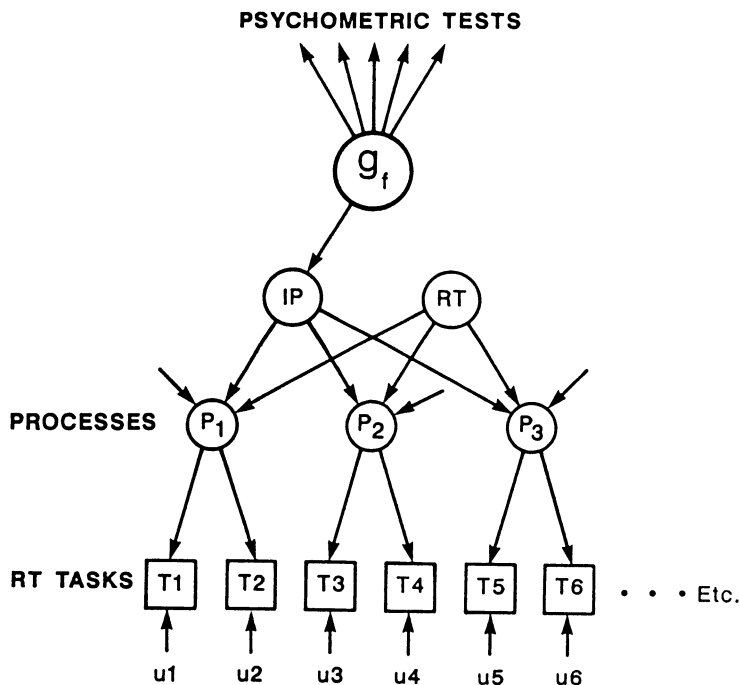
Still another hypothesis has been suggested by a factor analysis of RT data on a number of ECTs along with a diverse battery of psychometric tests obtained on 100 students tested in my chronometric lab (Kranzler & Jensen, 1991; see also Carroll, 1991). But first, it is important to know the leading alternative to this hypothesis, which has been most clearly enunciated by Detterman (1987). As shown in Figure 2, the RTs of a number of distinct ECTs representing different processes ( $P$ ) are each correlated with psychomet-



**Figure 2.** Representation of the factor structure of RT measures on a battery of diverse ECTs, in which some unspecified number information processes (P) tapped by the ECTs accounts for all of the variance in psychometric  $g$ .

ric  $g$ . If each process is independently correlated, say, .30 with  $g$ , then each one accounts for .09 of the variance in  $g$  and it would take (on average) about 11 such processes to account for all of the  $g$  variance. The problem is, when we actually include more and more different ECTs in a multiple regression to predict psychometric  $g$ , the squared multiple correlation ( $R^2$ ) rapidly approaches some asymptotic value that falls closer to .50 than to 1.00, even with correction for attenuation. Another problem with Detterman's hypothesis is that, unless we can actually find a number (*any* number) of ECTs that in combination can account for *all* of  $g$ , one could always argue that the *right* ECTs had not been tried or that some of the crucial cognitive processes had not yet been discovered.

These apparent problems, however, might simply evaporate if the hypothesis suggested by a factor analysis of the Kranzler and Jensen (1991) data (see Carroll, 1991) becomes well established by further evidence. The main features of the factor structure of these data are depicted in Figure 3. The total RT variance of the battery of ECT variables splits into two nearly equal parts when the ECTs are factor analyzed in conjunction with a battery of standard psychome-



**Figure 3.** Representation of the factor structure of RT measures on various ECTs in which only part of the RT variance is associated with information processes (IP) and part of it is due to noncognitive factors. Only the cognitive or information processing (IP) part of the total RT variance is related to psychometric  $g$ .

tric tests. Nearly half of the RT variance on the various ECTs (information processing speed, or IP in Figure 3) is found on the factor that is clearly identified as psychometric  $g$ , while the remaining RT variance is located on a separate group factor (RT in Figure 3), which might be called noncognitive RT, or certainly non- $g$  RT. The more complex ECTs have relatively higher loadings on the information processing (IP) component of RT and hence also on  $g$ , while the simpler ECTs have relatively larger loadings on the noncognitive RT factor, on which the psychometric tests have near-zero factor loadings. It is not known for certain what the noncognitive component of RT consists of; most of it is probably variance in the purely sensorimotor aspects of RT performance. This may also account for the generally higher  $g$  loading—about  $-.50$ —of *inspection time* (i.e., the speed of making a simple visual or auditory discrimination, which involves no motor component) than of RT based on any single ECT. (For a meta-analysis of research on inspection time correlations with  $g$ , see Kranzler & Jensen, 1989.) If this finding holds up in future studies, it may

be the case that we are already accounting for nearly all of the true  $g$  variance in terms of the speed of information processing component of RT measured on only a small number (8 in the Kranzler and Jensen study) of ECTs that involve several distinct information processes (in this study, stimulus apprehension, choice, discrimination, retrieval of information from STM, and retrieval from LTM). The  $g$  loadings of some of these RT variables are as large as the  $g$  loadings of some of the standard psychometric tests. Theoretically, if it were possible to rid the RT measurements entirely of their noncognitive variance, it should be possible to measure  $g$  solely with the RTs obtained from a small battery of ECTs just as well as by means of a large battery of psychometric tests that sample subjects' repertoire of past-acquired knowledge and complex reasoning and problem-solving skills. I expect that eventually we will be able to assess  $g$  directly from measurements of neural activity in the brain.

The possibility of such measurements would be a boon to those who wish to study secular changes in the overall level of general ability in the population. Measurements derived from ordinary psychometric tests are more or less context bound, hence scores are influenced by time and place. The information subtest of the Wechsler Adult Intelligence Scale, for example, is quite highly  $g$  loaded in the test's standardization population, yet some of the greatest intellects of the past—Plato and Archimedes, for instance—could not possibly give correct answers to more than three of the information items, which is an imbecile level of performance in the present standardization sample. Although the variance in psychometric test scores remains pretty much the same across time, and certain population groups seem to remain in the same *relative* positions, the overall central tendency of the score distribution may shift rather markedly over a period of two or three decades (Flynn, 1984, 1987).

Psychometric measurements are something like measuring the height of people by the length of the shadow they cast when standing in the sunlight. If all of the people's shadows are measured at the same time of day and at the same location on the earth, the measurements will be perfectly correlated with the direct measurements obtained in the usual way with a yardstick or tape measure, which of course would remain the same regardless of time and place, unless there were a true change in people's height. If we measure people's shadows at different times or locations, however, we could not tell if heights have really changed, unless we knew precisely how to take account of time and place in making the measurements. Or, still better, we could measure height directly with a ruler. Similarly, if we observe secular shifts in the overall distribution of our psychometric measurements in the population, we have no way of knowing to what extent the shift reflects some change in the biological anlage of ability, and to what extent it is due to some other type of effect, such as people having learned the particular test items, or acquired relevant information, or practiced similar cognitive skills. Improved nutrition of the population might be the cause of change in the one case, improvements in

education in the other. A possible solution to this problem would be to develop multiple regression equations that would *anchor* the psychometric test scores (or derived factor scores) to RTs on certain ECTs and to neurophysiological measurements afforded by evoked potentials, neural conduction velocity, the metabolic rate of cortical glucose, and the like—variables already found to be correlated with psychometric *g*. Such anchored scores would greatly aid analysis of the nature and causes of the secular changes in the overall distribution of scores on conventional mental tests in the population (Jensen, 1991).

## A THEORETICAL FORMULATION OF *G*

### Why Is Speed of Information Processing so Important?

The answer to this question rests on two empirically well-established facts: (a) the limited capacity of working memory (WM), and (b) the rapid loss of information in WM. Most probably these two facts are causally related; WM has limited capacity because of the rapid loss of information in WM. WM has been referred to as the “scratch pad” of the mind. Its functions consist of encoding incoming information, manipulating or transforming it as the task requires, rehearsing it for consolidation in long-term memory (LTM), and retrieving certain information stored in LTM demanded by the task at hand. It must perform any one or a combination of these functions before the neural traces of the recently received information have decayed beyond retrieval. Otherwise there is a loss of information, a “breakdown” in processing, and the input of information must be repeated if the problem is to be solved. Hence we must write down overly long phone numbers and solve complicated arithmetic problems with paper and pencil, because the amount of information involved and the number of mental operations that must be performed exceed the capacity of our working memory. Faster speed of information processing is advantageous because more information can be processed before it decays beyond retrieval. Some problems can be solved only by manipulating a number of items of information more or less simultaneously, so that if one item is lost, the problem cannot be solved or the necessary “insight” needed to achieve the solution cannot occur. Hence greater speed of processing information is a distinct advantage in any intellectually demanding pursuit.

Speed of processing is not necessarily related to the speed of selecting the correct answers in a multiple-choice test or even to the speed of solving complex problems, because in such cases there are differences in the depth and thoroughness of processing, which may take place rapidly but also extensively, thereby consuming more total time than a faster but more superficially derived response. High *g* individuals, therefore, usually display fast RTs to ECTs that

make minimal demands on the capacity of WM, and, at the other extreme, they can learn especially complex subjects, solve complex problems, and perform other complex mental feats that are beyond average and low-*g* persons, regardless of the amount of time allowed. In tasks of intermediate complexity, high-*g* persons usually process problems in greater depth (and hence have more correct solutions), but the solution to many such problems can be reached also with more superficial processing, though with greater risk of error, and so the average solution time per problem will not be highly correlated with RT to relatively simple ECTs or with *g*. Yet the *average amount of time* that it takes a group of subjects to solve each of a number of problems of varying complexity (e.g., the items of Raven's Matrices), given without time limit, is almost perfectly correlated with the *difficulty* of the problems, as indexed by the percentage of persons who fail to get the correct solution. This indicates the importance of speed in problem solving, even when speed is not ostensibly a requirement of the task, which is given with explicit instructions to take as much time as needed to attempt all the problems.

Imposing a strain (just short of the point of a "breakdown" of information processing) on WM capacity in ECT tasks (in which performance is measured by RT) rank orders subjects differently from the rank order of their RTs derived from ECTs that scarcely tax WM. Also, RT is more highly correlated with *g* when WM is taxed. It is necessary, therefore, to take WM capacity into account in our theory of *g*. At this point, a formulation of WM capacity by psychologists in Erlangen, Germany, which has some empirical support, seems to fill the bill (Lehrl & Fischer, 1988). In their formulation, the capacity (*C*) of WM is a function of the speed (*S*) of processing and the duration time (*D*) of information in STM, absent rehearsal. If amount of information is measured in bits (i.e., the binary logarithm of the number of choices or response alternatives), then  $C \text{ bits} = S \text{ bits/sec} \times D \text{ sec}$ . The Erlangen psychologists have empirically obtained estimates of these parameters in average adults, approximately, of  $S = 15 \text{ bits/sec}$ ,  $D = 5 \text{ to } 6 \text{ sec}$ , and  $C = 80 \text{ bits}$ . Assuming positive (but not perfect) correlations among *S*, *D*, and *g*, the measure of *C* theoretically should be more highly correlated with *g* than is RT or processing speed alone. Studies by the Erlangen group bear this out. Their measure of *C*, for example, correlated +.67 and +.88 with scores on a vocabulary test (a highly loaded variable) in two samples of adults, with *N*s of 672 and 66, respectively (Lehrl & Fischer, 1988).

In addition to WM capacity, formulated as  $C = S \times D$ , we also must take into account *oscillation* in speed of processing, indexed by SDRT. This is because SDRT, although highly correlated with RT or processing speed, is correlated with *g* independently of RT (Jensen, 1992e). The behavioral manifestation of oscillation is an empirical fact, but its causal mechanism is speculative at present. It most likely has some neurophysiological basis. For instance, we



know that neurons are periodically excitatory and refractory, and that large numbers of neurons may show synchrony in their oscillation in excitatory potential, which may be detected by electroencephalography. This could be the basis of the overt oscillation we see in RT measured as SDRT.

According to this theory, then, there are three properties of the brain that constitute the physiological basis of  $g$ , the general factor of mental abilities: (a) the *speed of neural conduction* (including synaptic delay) in the brain, (b) the rate of *oscillation* of excitatory potential in individual neurons and groups of neurons acting in phase, and (c) the *duration* (or conversely, the rate of decay) of the traces of recently input information in neural assemblies. Accordingly, a higher level of  $g$  is the result of faster neural conduction (NCV), a faster rate of oscillation, and a slower rate of decay of neural traces. While the evidence from RT studies indicates that *speed* and *oscillation*, though highly correlated, are also each independently correlated with  $g$ , suggesting that they are due to different properties of the nervous system, the relation between oscillation and the rate of decay of neural traces is more speculative. The decay of information in neural assemblies could be merely a product of oscillation. Oscillation may be thought of as neural “noise” in the transmission of information, which would reduce the overall efficiency of information processing and impair the capacity of WM.

Although oscillation of excitatory potential is a property of every nervous system, one might wonder why a rapid rate of oscillation is more favorable to  $g$  than a slower rate. If we think of oscillation as a neuronal “shutter,” analogous to the shutter of a camera, and if the “open” and “shut” phases of the shutter are rapid (i.e., of short duration), then little moment-to-moment detail will be lost, or shut out, from the continuous input of stimuli and the chaining of operations while processing information in WM. In the RT paradigm, for instance, if the onset of the reaction stimulus occurs during the subthreshold, or “off”, phase of neural oscillation, the signal will take longer for processing, which cannot be completed until the “on” phase occurs. Individuals with consistently more rapid oscillation of the “off” and “on” phases, therefore, will show less variability in RT over a number of trials. It is this variability that is negatively correlated with  $g$ .

### DISCLAIMER

The proposed theory of  $g$  is *not* a theory of individual differences in achievement or success in life. Although it is certainly true that  $g$  is related to certain types of achievement and to some criteria of success, it is but one, albeit an often important one, of the many different elements involved in these complex outcomes. It is granted that no conscious, voluntary behavior, including any act

involving mental ability, ever occurs in isolation, but always issues from a matrix of experience and knowledge, interests, motivation, values, and personality variables, as well as specific contextual or situational influences. It is granted also that, provided the level of  $g$  exceeds the level necessary for acquiring the knowledge and technical skills required for the person's particular pursuit, outstanding achievement depends on other ingredients more than on  $g$  per se, such as the development of specialized abilities, assiduous practice and the automatization of essential subskills, unflagging motivation, persistence in the face of difficulty, self-confidence, and a negligible fear of failure.

Nevertheless, the general factor of mental ability,  $g$ , can be distilled from this seeming welter of variables. The correlations of  $g$  with a host of "real-life" variables that, throughout the history of civilization have been regarded as important, not only to the individual but to society as a whole, make it probably the most significant factor of the human condition.

The long-sought explanation of  $g$  must eventuate as a specialized aspect of a theory of the human brain—its neurological structure, its physiology, its evolution, its ontogeny, and the genetic mechanisms involved in its variation. Pursuing this most fundamental goal should not, of course, preclude studies and theories of the multifarious manifestations of  $g$  in human behavior: (a) its interaction with other behavioral traits and the many environmental, experiential, and educational variables that may influence the expression of  $g$ ; (b) the study of various group differences in  $g$  and examination of the sociological, educational, and economic consequences of the wide range of variation of  $g$  in the population; (c) the development of cost-efficient tests of  $g$  for practical applications; (d) the investigation of other well-recognized mental ability factors independent of  $g$ ; and (e) the discovery of further authentic ability factors that are uncorrelated with presently established factors.

## REFERENCES

- Carroll, J. B. (1991). No demonstration that  $g$  is not unitary, but there's more to the story: Comment on Kranzler and Jensen. *Intelligence*, *15*, 423–436.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Detterman, D. K. (1987). What does reaction time tell us about intelligence? In P. A. Vernon (Ed.), *Speed of information processing and intelligence*. Norwood, NJ: Ablex.
- Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, *95*, 29–51.
- Flynn, J. R. (1987). The ontology of intelligence. In J. Forge (Ed.), *Measurement, realism and objectivity*. Norwell, MA: D. Reidel.
- Haier, R. J., Siegel, B. V., Nuechterlein, K. H., Hazlett, E., Wu, J. C., Park, J., Browning, J. L., & Buchsbaum, M. S. (1988). Cortical glucose metabolic rate

- correlates of abstract reasoning and attention studied with positron emission tomography. *Intelligence*, 12, 199–217.
- Jensen, A. R. (1980a). *Bias in mental testing*. New York: Free Press.
- Jensen, A. R. (1980b). Uses of sibling data in educational and psychological research. *American Educational Research Journal*, 17, 153–170.
- Jensen, A. R. (1982). Reaction time and psychometric *g*. In H. J. Eysenck (Ed.), *A model for intelligence*. New York: Springer-Verlag.
- Jensen, A. R. (1983). The effects of inbreeding on mental ability factors. *Personality and Individual Differences*, 4, 71–87.
- Jensen, A. R. (1987a). Psychometric *g* as a focus of concerted research effort. *Intelligence*, 11, 193–198.
- Jensen, A. R. (1987b). The *g* beyond factor analysis. In J. C. Conoley, J. A. Glover, R. R. Ronning, & J. C. Witt (Eds.), *The influence of cognitive psychology on testing and measurement*. Hillsdale, NJ: Erlbaum.
- Jensen, A. R. (1987c). Individual differences in the Hick paradigm. In P. A. Vernon (Ed.), *Speed of information processing and intelligence*. Norwood, NJ: Ablex.
- Jensen, A. R. (1988). The relationship between learning and intelligence. *Learning and Individual Differences*, 1, 37–62.
- Jensen, A. R. (1991). Speed of elementary cognitive processes: A chronometric anchor for psychometric tests of *g*. *Psychological Test Bulletin*, 4, 59–70.
- Jensen, A. R. (1992a). Comment: Vehicles of *g*. *Psychological Science*, 3, 275–278.
- Jensen, A. R. (1992b). Understanding *g* in terms of information processing. *Educational Psychology Review*, 4, 271–308.
- Jensen, A. R. (1992c). The importance of intraindividual variability in reaction time. *Personality and Individual Differences*, 13, 869–882.
- Jensen, A. R. (1993a). Psychometric *g* and achievement. In B. R. Gifford (Ed.), *Current views on testing: A policy perspective*. Boston: Kluwer.
- Jensen, A. R. (1993b). Why is reaction time correlated with psychometric *g*? *Current Directions in Psychological Science*, 2, 53–56.
- Jensen, A. R., & Reed, T. E. (1990). Simple reaction time as a suppressor variable in the chronometric study of intelligence. *Intelligence*, 14, 375–388.
- Jensen, A. R., & Sinha, S. N. (1993). Physical correlates of human intelligence. In P. A. Vernon (Ed.), *Biological approaches to the study of human intelligence*. Norwood, NJ: Ablex.
- Jensen, A. R., & Weng, L.-J. (in press). What is a good *g*? *Intelligence*.
- Kranzler, J. H. (1992). A test of Larson and Alderton's (1990) Worst Performance rule of reaction time variability. *Personality and Individual Differences*, 13, 255–262.
- Kranzler, J. H., & Jensen, A. R. (1989). Inspection time and intelligence: A meta-analysis. *Intelligence*, 13, 329–347.
- Kranzler, J. H., & Jensen, A. R. (1991). The nature of psychometric *g*: Unitary process or a number of independent processes? *Intelligence*, 15, 397–422.
- Larson, G. E., & Alderton, D. L. (1990). Reaction time variability and intelligence: A "worst performance" analysis of individual differences. *Intelligence*, 14, 309–325.
- Lehrl, S., & Fischer, B. (1988). The basic parameters of human information processing: Their role in the determination of intelligence. *Personality and Individual Differences*, 9, 883–896.
- Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor of job performance.

- Current Directions in Psychological Science*, 1, 86–89.
- Reed, T. E. (1988). A neurophysiological basis for the heritability of intelligence. In H. J. Jerison & I. Jerison (Eds.), *Intelligence and evolutionary biology*. Berlin: Springer-Verlag.
- Reed, T. E., & Jensen, A. R. (1992). Conduction velocity in a brain nerve pathway of normal adults correlates with intelligence level. *Intelligence*, 16, 259–278.
- Schmidt, F. L., Ones, D. S., & Hunter, J. E. (1992). Personnel selection. *Annual Review of Psychology*, 43, 627–670.
- Stankov, L. (1988). Single tests, competing tasks and their relationship to the broad factors of intelligence. *Personality and Individual Differences*, 9, 25–33.
- Sternberg, R. J., & Detterman, D. K. (Eds.). (1986). *What is intelligence?* Norwood, NJ: Ablex.
- Vernon, P. A. (1983). Speed of information processing and general intelligence. *Intelligence*, 7, 53–70.