# Spearman's *g*: Links Between Psychometrics and Biology

ARTHUR R. JENSEN
*School of Education*
*University of California*
*Berkeley, California 94720*

Charles Spearman[1] discovered *g* almost 90 years ago and it has continued to be a major construct of differential psychology. Interest in theory and research on *g* has waxed and waned over the nine decades of its venerable history, but it is probably greater today than in any previous period.

In recent years, researchers have made important strides toward answering some of the most crucial questions and controversies regarding the nature of *g*. Our understanding is now well beyond the point where most theorists are satisfied with merely psychological explanations of one of psychology's major phenomena. The strictly psychometric aspects of *g* are now so well established empirically and so firmly grounded methodologically as to no longer command a further major research commitment at this level. As the current trend of theory and research on *g* advances closer to the interface of brain and behavior, it seems a safe prediction that the main line of progress toward our future understanding of *g* will be in the province of biology, particularly the brain sciences, but also genetics and evolutionary neurology.

## 1. *The Concept of Psychometric* g

Various tests of mental ability reflect rather different abilities, as shown by the fact that when scores on a number of such tests are obtained from a representative sample of the general population, the correlations between people's scores on the various tests are less than perfect, even when corrected for measurement error, or unreliability. Correlations among different mental tests generally range from about +.20 to +.80, depending mainly on how much the various tests vary in task complexity, diversity of information content, and required skills. Although the test intercorrelations are less than perfect, the theoretically intriguing point is that they are all positive and, when obtained in fairly large and unrestricted samples, are virtually all significantly greater than zero.

This empirical phenomenon, which is now one of the most solidly substantiated facts in psychology, can be interpreted to mean that all kinds of mental ability tests measure something in common. Spearman called this "something" the *general* factor, which he symbolized as *g*. Whether called "Spearman's *g*," "psychometric *g*," or just plain *g*, this construct refers to the component of individual differences variance that is common to all tests of mental ability. In addition to discovering the existence of *g*, Spearman invented a method by which the coefficient of correlation between a particular test and the *g* factor could be determined. (Such a correlation is termed the test's *g loading.*)

The *g* factor is the sine qua non of all "IQ" tests, no matter what other sources of variance such tests may measure. The *g* factor is commonly thought of as "general intelligence." As a short definition of *g*, however, "general mental ability" is preferable to "general intelligence." For reasons spelled out elsewhere,[2,3] the persistently indefinable and emotive word "intelligence" is best avoided in the discussion of human abilities. "Intelligence" has so many different meanings, especially among psychologists, and carries so much excess connotative baggage as to be a hindrance to serious discussion.

### 1.1. *The Psychometric Nature of* g

Before getting to the main topic, several of the most germane and well-established facts about psychometric *g*, which were obtained strictly from factor analyses of a wide variety of conventional tests of mental abilities, should be reviewed. "Mental" here only means that a negligible part of the individual differences' variance in performance is attributable to sensory-motor functions, while "ability" is defined as an intentional response to some situation that can be classified or scaled in terms of an objective criterion of the "goodness" of the subject's performance (e.g., correct or incorrect response, response time).

**1.1.1.** In over 80 years of research on human abilities, no cognitive task has been consistently found to be not positively correlated with any other cognitive tasks, or not show a positive *g* loading when factor analyzed among a fairly large battery of diverse tests of mental abilities, administered to a representative sample of the general population. Consequently, the main tenet of Spearman's *g* theory has never been refuted, since every reliable indicator of mental ability has some positive *g* loading.

**1.1.2.** The *g* factor and other factors derived from factor analysis do not represent the operating principles, processes, mechanisms, or structures of the mind or the brain. In principle, these properties of the brain could be discovered by studying only one individual. Factors, on the other hand, depend entirely on individual differences, or variance; on a number of distinct variables; and on all the correlations between them. This necessitates studying more than one individual. Within a given domain of variables, factor analysis can represent the "structure" of individual differences. The term "structure" here refers to the representation of a pattern of correlations in terms of a mathematical or geometric structure, or model, with no direct implications regarding any physical or anatomic structures. However, certain results of factor analysis, such as the discovery of *g*, may lead to questions that can be answered only in terms of physical structures and processes in the brain.

**1.1.3.** A *g* factor can be extracted from a correlation matrix by several different methods. It may surprise many that *g* obtained as either the first principal component or principal factor is a conceptually inappropriate factor model for extracting *g*. In theory, using principal component or principal factor analysis can yield misleading results. For example, a simulated correlation matrix that has been expressly devised to exclude a general factor can appear to have a very substantial *g* factor. However, if a matrix actually contains a general factor that accounts for a sizable proportion of its total common factor variance, both of these methods will yield reasonably accurate estimates of the *g* loadings of the variables in the matrix. Since real ability test matrices nearly always do contain a large general factor, the vast majority of past studies using principal components and principal factors have not given a misleading picture. Furthermore, past results are highly congruent with the *g* factors extracted by means of the logically most appropriate and statistically sophisticated methods available, methods which have been tested on simulated variables whose true factor structure is known exactly.

**1.1.4.** The *g* factor is remarkably stable across different test batteries, and the stability increases as a function of both the number and diversity of tests that are entered into the factor analysis. Tests with very high *g* loadings within one particular battery of diverse tests remain highly *g*-loaded when they are included in factor analyses of other diverse batteries. In a study of the stability of *g* by Robert L. Thorndike,[4] for example, there was an average correlation of +.85 between the *g* loadings of 17 different "probe" tests when included one at a time in the separate factor analyses of 8 distinct batteries, each composed of six highly diverse tests. In other words, each of the "probe" tests' *g* loadings maintained about the same relative position among the *g* loadings of all of the other 16 "probe" tests, regardless of the particular battery of tests in which it was factor analyzed. The point is that the *g* factor (and ipso facto any particular test's *g* loading) does not fluctuate capriciously from one collection of tests to another. If a battery of tests is quite diverse in sampling the domain of abilities and is fairly large (i.e., ten or more tests), its *g* factor will be highly stable, that is, highly correlated with the *g* of any other test batteries having these characteristics. For example, the six Verbal subtests of the Wechsler Intelligence Scale for Children (WISC) look very different from the six Performance subtests, but the *g* extracted from just the Verbal subtests is correlated about .80 with the *g* extracted from just the Performance subtests.

**1.1.5.** In a large and highly diverse collection of mental tests, the various tests' *g* loadings are a perfectly continuous variable, ranging from slightly greater than zero to near the ceiling of test reliability. Ability tests do not fall into discrete categories with respect to their loadings on *g*. However, it is possible to characterize the types of tests that have high or low *g* loadings. As noted by Spearman, the most highly *g*-loaded tests involve complex cognitive operations, such as "the eduction of relations and correlates," or inductive and deductive reasoning, and "abstraction." Tests with low *g* loadings generally involve less complex cognitive processes, such as simple sensory discriminations, reaction times to simple stimuli, and rote memory. In the WISC, for example, the most highly *g*-loaded subtests are Vocabulary and Block Design;

the least *g*-loaded are Coding and Digit Span. A test's *g* loading is seldom predictable from any such specific features as its sensory modality or knowledge content per se, but is much more predictable on the basis of the subjectively perceived complexity of the mental operations required for passing performance. For example, the backward digit span test (i.e., recalling a series of digits in reverse order) involves slightly more complex mental operations than forward digit span, and backward digit span has almost twice the *g* loading of forward digit span. Both of these relatively simple tests, however, have rather low *g* loadings compared with the more complex subtests of the WISC.

**1.1.6.** Psychometric *g* cannot be described in terms of the superficial characteristics of tests—the specific knowledge, skills, or problem solving strategies they may involve. The *g* factor derives much of its interest and importance from the fact that it is *not* a measure of specific knowledge, skills, or strategies for problem solving. Broadly speaking, *g* reflects individual differences in information *processing* per se. The knowledge and skills aspect of mental test performance is merely a *vehicle* for the measurement of *g*, which reflects the overall capacity and efficiency of the brain processes by which knowledge and skills are acquired and used. Hence we cannot begin to fathom the causal underpinning of *g* by merely examining the psychometric tests themselves.

Because *g* emerges from the analysis of correlated individual differences on a large number of very diverse tests, the variance attributable to specific characteristics and task demands of the particular tests are in effect averaged out. This is also true for the common characteristics of certain classes of tests, such as verbal, numerical, spatial, memory, and the like, whose intercorrelations (independent of *g*) form *group* factors. Ideally, the *g* factor may be thought of as a distillate of the common source of individual differences in all of the tests, completely stripped of their distinctive features of information content, skill, strategy, and the like. In this sense, *g* can be roughly likened to the Central Processing Unit (CPU) of a computer, and like the computer's CPU, *g* seems to reflect individual differences in the brain's "hardware" more than in its "software."

**1.1.7.** Individual differences in human learning abilities also reflect *g*—the same *g* derived from psychometric tests. That is, the general factor in a wide variety of learning tasks is found to be the same factor as psychometric *g*. There seems to be no general learning ability factor independent of *g*, but there are many group factors and a lot of specificity in the wide variety of learning tasks that have been studied. The research on the relationship between learning and *g* has been reviewed elsewhere.[5]

**1.1.8.** There is now a vast literature showing that *g* is the chief "active ingredient" in the practical predictive validity of psychometric tests for such criteria as educational achievement, occupational level, job performance, and success in armed forces training programs.[4,6–9] What this implies, of course, is that many "real life" tasks, especially educational and occupational demands, are *g*-loaded to varying degrees and tend to discriminate among people on this basis.

## 2. *Correlates of* g *outside the Psychometric Domain*

One of the most remarkable properties of *g*, in contrast to other psychometric factors independent of *g*, is that *g* shows correlations with a host of variables whose measurement and conceptual basis are completely separate from psychometrics and factor analysis.[10,11] This fact alone proves that *g* is not merely some kind of artifact of psychometric tests or the mathematical machinations of factor analysis.

The degree to which various psychometric tests are *g* loaded is highly related to their degree of correlation with such nonpsychometric variables as the *heritability* of individual differences in test scores, the *spouse correlations* and various *genetic kinship correlations* in the test scores, and the *effects of inbreeding* (and its counterpart, *heterosis*) on test scores. The fact that the degree of inbreeding depression of scores on various tests (as observed in the offspring of cousin matings) is highly related to the size of the tests' *g* loadings indicates that *g* is a genetically dominant trait, and genetic dominance arises from natural selection in the course of evolution.[12] The presence of genetic dominance indicates a fitness character in the Darwinian sense. Thus it appears that *g* is deeply rooted in biology.

Also, certain features of the electrical activity of the brain, such as the latency of the *average evoked potential* (AEP), are correlated with various psychometric tests. The degree of a test's correlation with the AEP is directly related to the size of the test's *g* loading. It is especially noteworthy that no other sources of variance in psychometric tests show any correlation with the AEP.[13]

Many other physical variables (i.e., anatomical, physiological, serological, and biochemical) are correlated with *g*. For most of these correlations, however, the chain of causality is so indirect, and so obscure at present, as to be of little help in understanding the nature of *g*. Thus, although very real in a statistical sense, these correlations are mysterious and must await future explanation. Without a doubt, however, they show that some part of the population variance on conventional tests of mental ability reflects certain latent traits, primarily *g*, that are profoundly enmeshed with many organismic variables in complex ways.

### 2.1. *Between-Families and Within-Families Correlations*

In a comprehensive review of the evidence relating to these various physical correlates of human intelligence,[14] an important distinction is made between two main types of correlation between *g* and physical variables: (1) a correlation that is found only *between* families but not *within* families, and (2) a correlation that is found *within* as well as between-families (where "family" means only full siblings). As explained elsewhere,[15] this distinction is theoretically crucial in limiting the kinds of hypotheses that could explain each type of correlation. Although both types of correlation may be equal in size and statistical significance, their causal underpinnings are importantly different. The first type of correlation (i.e., between families only) rules out genetic pleiotropy, that is, two or more distinct phenotypic characteristics whose variance is affected by the same gene (or set of genes). The correlation of about +.20 between height and *g*, for example, is only a *between*-families correlation; there is no correlation between height and *g* among full siblings. The between-families correlation in this case seems to have come about as the result of the common assortment of genes for height and for *g*, since both of these variables partly influence mate selection in our culture. Other things being equal, high *g* women tend to marry tall men (and vice versa), and since both traits are highly heritable, their offspring (on average) tend to be both taller and higher in *g* than average. On the other hand, myopia, which is correlated about +.25 with *g*, is a *within*-families correlation, suggesting that it is a pleiotropic correlation, in which one or more of the genes involved in myopia also have the effect of enhancing *g*. The sibling with higher *g* is more likely to be myopic.[16] So myopia probably has some pleiotropic relation to brain functions, but the neurology of the causal connection and how it came about in the course of human evolution are completely unknown.

Head size and brain size are correlated with *g* independently of body size. From a meta-analysis of all the published studies, the estimated correlation between brain size and IQ (or other highly *g* loaded indices) is +.30.[14] A recent study[17] of 40 healthy college students measured brain size quite accurately *in vivo* by means of magnetic resonance imaging (MRI) and correlated these measurements with the students' Wechsler IQs. From the results, the authors inferred a brain size × IQ correlation of +.35 adjusted for body height and weight, in the general population (with mean IQ = 100, SD = 15). Neurological theories of *g*, of course, will have to explain the brain-size × *g* correlation, but the explanation will be neurologically informative only if the correlation exists *within* families. This crucial point has not been firmly established, although it would be most surprising if the correlation between brain size and IQ were found to exist only *between* but not *within* families, like the correlation between height and IQ.

Speculative theories of the brain-size × IQ correlation have invoked variables such as the number and amount of branching of brain cells, the number of synaptic connections, and the number of glial "support" cells, or some combination of these. One of the unsolved mysteries is that males and females differ in brain size, even after adjustment for the average sex difference in body height and weight, despite the failure to find any good psychometric evidence for an average sex difference in *g*. Yet *within* groups of males and females separately, brain size is positively correlated with *g*.

**2.1.2.** Other physical variables that seem to correlate with *g*, but for which the *causal* connection is either largely unknown or hard to hypothesize, are certain blood antigens, serum uric acid level, vital capacity, facial features, basal metabolic rate (in children), asthma and other allergies, presence or absence and size of the masa intermedia (which connects the two halves of the thalamus), ability to taste the synthetic chemical phenylthiocarbamide (PTC), and tongue-curling ability (which is controlled by a single gene).

### 2.2. *Reaction Time (RT) on Elementary Cognitive Tasks (ECTs)*

This is a class of variables whose relation to *g* is now well established and whose properties are much better suited than those of conventional psychometric instruments for testing hypotheses concerning the neural basis of individual differences. RT has the advantage of being a true ratio scale and does not depend on a norm or reference group for its interpretation, as do conventional psychometric tests. ECTs also permit the measurement of individual differences while minimizing variance attributable to specific knowledge and acquired intellectual skills and problem-solving strategies. ECTs are devised to reflect individual differences in information *processes* rather than in the specific *content* of information.

ECTs are such simple mental tasks that virtually everyone can perform them correctly and easily, making RT the only reliable source of variance. In our research on the correlations of RT with psychometric *g*, the ECTs are so easy that even for the most complex tasks the RTs average less than one second and response errors average less than 5%. Obviously, such tasks could scarcely involve what one would think of as cogitation.

Typical ECTs consist of various forms of simple, choice, and discrimination RT (e.g., Hick and odd-man-out paradigms); *visual scanning* of a set of one to seven digits and indicating the presence or absence of a pre-cued "target" digit; *memory scanning* of a set of one to seven digits held in short-term memory and indicating the presence or absence of a post-cued "target" digit (i.e., the Sternberg paradigm); *retrieval from long-term memory* of highly over-learned lexical information, using variations of the Posner paradigm in which the subject responds "same" or "different" to pairs of highly familiar words that are either synonyms or antonyms (e.g., "little-small," "hot-cold"); a *semantic verification* test in which two or three letters (e.g., AB or ABC) are presented for three seconds, followed by a statement (e.g., "B after A" or "A before B") to which the subject responds either "True" or "False"; and *inspection time* (IT) in which two vertical lines, one twice the length of the other, are presented tachistoscopically side-by-side, followed by a "masking" stimulus, and the subject then has to indicate (without time pressure) whether the longer line appeared on the right or the left side—the subject's IT is the length of the interval between the onset of the pair of lines and the onset of the "mask" at which the subject can respond with 95% accuracy on 20 consecutive trials. All of these paradigms yield highly reliable measures of RT (or IT), and these measures are negatively correlated with scores on conventional psychometric tests. It is an important fact that the magnitude of these correlations is directly related to the degree of the psychometric tests' *g* loadings rather than to any specific type of information content or to whether the psychometric tests were administered under speeded or nonspeeded conditions.[18] (For more detailed descriptions of these ECTs and evidence of their correlations with *g*, readers may consult References 2, 11, 19, 20, 21, 22, 23, and 24).

### 2.3. *Some Salient Facts and Questions Arising from RT Studies*

During the past decade a great many studies have been done on the relationship between RT variables, as measured in various ECTs, and psychometric *g*. In most studies *g* is measured by a single test that is highly *g*-loaded, such as Raven's Progressive Matrices. Space does not allow description of particular experiments, but it is possible to summarize briefly some of the typical findings and the key questions they have prompted.

In order to understand the results of the studies to be discussed, one needs some idea of how the data were obtained. ECTs are performed on a response console that permits the subject to respond to the reaction stimulus (RS) by means of pushbuttons, operated by just the index finger of the subject's preferred hand. To begin a trial the subject depresses a central "home" button. Then a 0.5-second auditory preparatory signal ("beep") sounds, and after a random interval of one to four seconds the reaction stimulus (RS) appears. The subject is instructed to release the home button and press the response button as quickly and accurately as possible. The number of response buttons used in various ECTs varies anywhere from one to eight, all equidistant (six inches) from the home button. The response buttons are translucent and each one contains a light bulb, so that in some ECTs (e.g., simple, choice, and discrimination RT) the RS consists simply of one (or more) of the response buttons lighting up conspicuously. In ECTs involving a binary choice the two response buttons are appropriately labeled (e.g., "Same-Different," True-False," or "Yes-No"). A computer automatically administers the ECT and records the subject's performance. The three most important variables measured (in milliseconds) by this setup are operationally defined as follows.

*Reaction Time* (RT) is the interval between the onset of the RS and the subject's release of the home button. (This interval has also been termed "Decision Time," or DT, by some researchers, but DT seems an inappropriate term when the interval comprises sensory-motor time as well as the central processing time involved in making a decision. Only if the amount of time for the sensory-motor activity per se could be removed from this interval would DT be an accurate label. Otherwise, RT is the preferred term.) The subject's overall RT score is the median RT over *n*, the number of test trials on a given ECT.

*Movement Time* (MT) is the interval between the subject's releasing the home button and pressing the response button. The overall MT score is the median MT over *n* test trials.

*Intraindividual Variability in RT* (SDRT) is the standard deviation (SD) of the subject's RTS over *n* test trials.

**2.3.1.** *The Correlation between RT and Psychometric* g. In single ECTs, the zero-order correlations (Pearson *r*) between RT and *g* are typically in the

–.2 to –.4 range, seldom larger, although with proper corrections for attenuation and restriction of range, they may increase to about –.5. Investigators have referred to the "–.35 barrier," because that value is so typically near the upper limit of the most consistently replicable correlations (uncorrected) between *g* and RT obtained from any single ECT of the type previously described.

The *r* can be increased appreciably by combining the RTs from several different ECTs, either unit-weighted by simple addition or optimally weighted by multiple regression. Such correlations are generally in the –.4 to –.5 range, and with correction for attenuation and restriction of range they can reach about –.6.

Thus we have three well-established findings that together raise important theoretical questions: (1) Why is RT correlated with *g*? (2) Why is there a true "ceiling" on the correlation between RT and *g* for any single ECT or even any combination of ECTs? And (3) Why is the correlation between RTs and *g* increased by combining RTs from different ECTs?

**2.3.2.** *The Correlation between SDRT and* g. Intraindividual variability in RT, or SDRT, is not a measure of speed but of inconsistency of the subject's RTs across a number of trials. It is usually more highly correlated (negatively) with *g* than is RT, despite the fact that the reliability (split-half and test-retest) of SDRT is considerably lower than that of RT. The fact that combining both RT and SDRT (as *z* scores) yields a larger correlation with *g* than is found for either variable alone suggests, but does not prove, that RT and SDRT tap independent components of *g*. More telling is the fact that the zero-order correlation between RT and SDRT is high (+.6 to +.7), but it remains considerably less than perfect after correction for attenuation (about +.8), which means that individual differences in these two variables are not merely different manifestations of a single source of variance. This finding, too, calls for theoretical interpretation.

**2.3.3.** *MT Correlations.* MT is much faster than RT (except in the severely retarded) and, unlike RT, it varies remarkably little as a function of task complexity in different ECTs. MT is less consistently correlated with *g* than is RT. Correlations are typically in the –.10 to –.20 range, and in samples of college students are quite commonly close to zero. Mentally retarded adults, young children, and the elderly show higher correlations between MT and *g* as well as higher correlations between RT and MT. In young adults of average and superior ability, the correlations between RT and MT are only +.3 to +.4, and factor analyses of RTs and MTs measured on various ECTs show that RT and MT have their largest loadings on different factors. All these findings underscore the importance of measuring RT and MT separately. An amalgam of both variables often weakens correlations with *g*, and in some studies MT may even act as a suppressor variable for the correlation between RT and *g*, so that subtracting MT from RT increases the correlation of *g* with the difference, RT–MT, over that obtained for RT.[25] An obvious interpretation is that subtracting MT from RT removes a motor-speed component from

RT, leaving a purer measure of the information processes, which are related to *g*.

**2.3.4.** *RT-*g *Correlation as a Function of Task Complexity.* The correlation of RT with *g* increases as a function of independently rated judgments of the complexity of the ECT, but only up to a point, beyond which the correlation decreases. Hence the relationship resembles an inverted U. In young adults, the correlations peak on tasks for which the mean RT is between about 0.8 and 1.2 seconds. The RT-*g* correlation decreases rapidly for levels of task complexity that make for longer RTs. Also, response errors increase on more complex tasks and at some point the correlation between errors and *g* greatly exceeds the correlation between RT and *g*.

Our interpretation of this inverted U relationship of the correlation between RT and *g* as a function of task complexity involves three elements: (1) the more complex ECTs enlist a larger number of different elementary processes and, because *g* reflects all such processes, there are more processes in common (hence higher correlation) between the more complex ECTs and *g*. (2) More complex tasks require more time for information processing and hence put a strain on the subject's *working memory*, an aspect of short-term memory that is a processing system with the properties of limited capacity and rapid loss of recently input information. If processing speed loses the race with the rate of loss of the information in working memory, there is a "breakdown" in solving the "problem" posed by the ECT, and the subject's overt response will be at the chance level for errors. (3) In the more complex ECTs, subjects may resort to various strategies for getting around the limitations imposed by processing speed and the capacity of working memory. Subjects differ in the tendency to adopt strategies, as well as in the stage of practice at which they do so, and they may adopt different strategies of unequal effectiveness. These kinds of individual differences, at least as they are manifested in the realm of ECTs, apparently have little or no relation to *g* and are included in that part of the variance on tasks attributable to some combination of group factors and specificity. Hence their presence in the more complex tasks lessens the correlation of RT with *g*, while increasing the correlation of rate of response errors with *g*.

This reciprocal relation between RT and error rate, with respect to their correlation with *g* as a function of task complexity, was seen most strikingly in a study[20] in which a set of 14 ECTs (semantic verification tests) that differed in complexity was administered to college students and to third-grade school children. In the college sample, RTs were measured on each task, and these ranged, on average, from about .6 sec to 1.4 sec, with a 7% error rate. These tasks were relatively much more complex for the third-graders, who were administered the very same tasks, but with instructions to take as much time as they needed for every task. Even without time pressure they had a 17% error rate. The important point, however, is that there was a high correlation between the mean RTs of the college students on the 14 tasks and the mean error rates of the third-graders on the corresponding tasks. Thus in tasks of low complexity, individual differences in *g* are manifested in RT. As task complexity increases, however, relative to the subjects' speed of information processing and the capacity of working memory, *g* is decreasingly reflected by RT and increasingly reflected by response error rate.

It should also be noted that the relation between the RT-*g* correlation and task complexity is still problematic, in that there are numerous examples in which the predicted relationship is borne out only slightly or not at all, for reasons that so far are only speculative.[26] This problem most likely involves the uncertainty as to the position of the inverted U function with respect to the variable of task complexity, a position that varies in different groups according to their level of ability.

The nature of the differences between correct and erroneous RTs was illuminated at a physiological level in a study of the trial-by-trial correlation between the subjects' overt RTs and the latencies of the P300 cortical potential (EP) evoked by the reaction stimulus (RS) on the same trial.[27] The ECTs used in this study were three different binary choice RT tasks varying in the complexity of the hypothesized processes involved in the evaluation of the RS and administered under two conditions of speed/accuracy instructions. Correlations between the latencies of the P300 EP and the RTs on *correct* responses ranged from +.48 to +.66, but the correlations based on *both* correct and incorrect RTs were considerably lower (+.26 to +.61). More importantly from the standpoint of RT error theory, however, is that not only are the correct RTs relatively more closely coupled with their corresponding P300 EP latencies than are the incorrect RTs, but with remarkably few exceptions, the *correct* RTs are *slightly longer* than their corresponding EP latencies, while the *incorrect* RTs are almost invariably *much shorter* than their corresponding EP latencies.

These findings indicate that on the incorrect RT trials the overt reaction process was initiated long before the process associated with P300 was terminated. Kutas et al.[27] hypothesize a dual nature of RT as consisting of "stimulus evaluation" processes (reflected by the P300 EP) and the efferent processes involved in executing the overt response. As there is independent evidence that P300 reflects completion of the most central evaluative discrimination or decision aspect of information processing, it appears that RT response errors occur when the extent of information processing of the RS required for a correct response is for any reason incomplete.

**2.3.5.** *"Peripheral" and "Central" Components of RT.* RT reflects two conceptually distinct components that can be referred to as *peripheral* and *central.* The peripheral aspect includes time lag associated with sensory transduction of the RS plus the motor nerve conduction and muscle action involved in response execution. The central aspect of information processing presumably lies entirely within the cerebral hemispheres. This locus of information processing includes stimulus encoding and decision time (DT) and is primarily responsible for the correlation between RT and *g*. As noted in Section 2.3.3, simple RT has lower correlations with *g* than complex RTs (i.e., RTs in response to relatively more complex ECTs, such as choice or discrimination RT). We have hypothesized that this is due to the fact that a larger proportion of simple RT than of complex RT (which is always greater than simple RT) is peripheral.

Since variance in the peripheral component is unrelated to variance in the speed of information processing, it only attenuates the correlation between RT and *g*, just as would be the case if we added a random error component to RT. All forms of RT, of course, comprise both peripheral and central time increments, but the more complex RTs comprise larger increments of central processing time. This hypothesis receives support from our finding that the correlation between discrimination RT and *g* is increased by subtracting every subject's simple RT from the discrimination RT, so only the central component of discrimination RT remains to be correlated with *g*.[21] Thus the peripheral component of RT may act as a suppressor variable in the correlation between RT and *g*.

**2.3.6.** *Genetic Correlation between RT and* g. Three studies provide evidence that the relationship between RT and *g* is not only phenotypic but also genotypic, that is, they both reflect some common component of genetic variance. Intellectually gifted children have both higher *g* and faster RTs than their less gifted siblings, and the RT difference between the gifted and their siblings increases as a function of ECT complexity.[19] Hence this is a *within*-family correlation, which suggests pleiotropy, i.e., the same genes causing individual differences in both RT and *g*. A pleiotropic correlation between the cognitive component of RT and *g* would be explained most parsimoniously in terms of their sharing a common mechanism, namely, speed of information processing. Of course, the within-family correlation could also be caused by some unknown environmental factor that has an effect on both RT and *g*. However, there are two other studies that make genetic pleiotropy the much more compelling explanation. First of all, recall that the *g* loadings of various psychometric tests are directly related to the tests' heritability coefficients. Also, it has been found, using the MZ-DZ twin method for estimating heritability, that the size of the correlation between various measures of RT and *g* is a function of the heritabilities of the various RTs.[28] Another study, this one based on twins reared apart, also found RT measures of information processing to have substantial heritability.[29] Finally, by means of a sophisticated multivariate biometrical genetic analysis based on twin data, it was possible to determine the genotypic correlation between RT and *g*. The analysis showed a common genetic influence on both RT and *g*.[30] The authors concluded that their finding supports the hypothesis that speed-of-processing and IQ may share some common biological mechanism(s).

**2.3.7.** *Distinction between Process Differences and Individual Differences.* It is theoretically important to recognize that ECTs which reflect different cognitive processes do not necessarily interact with individual differences. In an analysis of variance of any study based on RTs derived from two or more different ECTs, there are *main effects* for mean differences between ECTs and between subjects, as well as an ECTs × subjects *interaction*. But even with highly significant main effects, the interaction may be negligible and nonsignificant. In other words, although different ECTs can be expressly devised to elicit RTs that clearly involve different elements of the information processing system, individual differences may remain invariant across the different ECTs. This finding indicates that one and the same cause of the individual differences acts in both of the different processes. A simple mechanical analogy would be two different machines that operate at different speeds and perform quite different jobs, but their different output rates are perfectly correlated because they are both connected by different sized cogwheels in a gear-train driven by a single constant-speed motor. Across different sets (analogous to individual differences) of these machines, the only source of (individual) differences in output rates would arise from consistent differences in the speed of the motor that drives each set. A striking example of this phenomenon in the realm of ECTs is seen in a comparison of the RTs obtained in visual search (VS) and memory search (MS) tasks. There are clear process differences in these tasks, since the role of short-term memory is nil in the VS task and is crucial in the MS task. The process difference is reflected in the significantly different average RTs obtained in these tasks. But individual differences in RTs on the VS and MS tasks are found to be perfectly correlated after correction for attenuation.[31] Process differences between certain other ECTs, however, do show interaction with individual differences, so the disattenuated correlation between the RTs on these tasks may be rather low, indicating that they tap independent sources of variance in addition to having some source of variance in common.

What all this means is that the sources of individual differences in various ECTs do not correspond directly to the different information processes that are hypothesized to be involved in the ECTs. Because individual differences do not necessarily coincide with different processes, but cut across different processes, it may be ultimately futile to try to map the basic dimensions of individual differences in terms of the various hypothesized processes of cognitive psychology. Similarly, *g* defies description in terms of the observable characteristics of psychometric tests. In this same connection, Willerman and Bailey offer the following conjecture: "Correlations between phenotypically different mental tests may arise, not because of any causal connections among the mental elements required for correct solutions or because of the physical sharing of neural tissue, but because each test in part requires the same 'qualities' of brain for successful performance. For example, the efficiency of neural conduction or the extent of neuronal arborization may be correlated in different parts of the brain because of a similar epigenetic matrix, not because of concurrent functional overlap."[32] In other words, correlation without direct functional relationship. They go on to point out certain functional abnormalities or damage to brain structures that severely impair a particular cognitive ability yet may not show even the slightest adverse effect on some other ability, even though individual differences in the two abilities are normally highly correlated. This relationship is analogous to the relationship between strength-of-grip of the left and the right hand, which are highly correlated in the population. A complete paralysis of one hand, however, may have no effect on the other. Such evidence clearly contradicts Thomson's "sampling" theory, which holds that the correlation between mental tests is a function of the number of neural elements, or bonds, that they involve in common.

## 3. *Recent Answers to Key Questions about the RT-*g *Correlation*

### 3.1. *Is Psychometric* g *a Unitary Process or Multiple Processes?*

This has been a contentious point throughout the history of *g*. There are two rival hypotheses.

**3.1.1.** The *unitary process* hypothesis holds that some single property of the central nervous system (CNS) in which there are individual differences, such as in the number of nerve cells, the amount of dendritic branching, the speed of neural conduction, metabolic efficiency, or richness of the capillary blood supply, etc., determines some part of the variance in the speed, efficiency, or accuracy of performance on all cognitive tasks, causing them all to be positively correlated and hence to yield a *g* factor.

Tasks differ in *g* loading, according to this hypothesis, for two main reasons:

(1) The more complex tasks or test items evoke more extensive neural activity for a longer period of time, thereby providing a larger and more reliable sampling of some unitary property of the CNS. This results in more complex tasks' being more highly correlated with one another and hence more highly *g*-loaded.

(2) Tasks or tests differ in the degree to which specific acquired knowledge, skills, and strategies affect performance, so tests differ in the proportions of their variance attributable to these experiential differences and hence differ in *g* loadings to the extent that the variance due to differences in experiential influence is independent of *g*.

It has been impossible to disprove the unitary hypothesis by any analytic techniques when their application is confined to scores on conventional complex mental tests from which *g* is ordinarily derived by means of factor analysis.

**3.1.2.** The *multiple processes* hypothesis holds that *g* is attributable to a number of independent, or uncorrelated, processes, in which there are individual differences. The all-positive correlations among the relatively complex mental tests that give rise to *g* result from the fact that a number of independent elemental processes are involved in any particular test, so various tests all have some processes in common and hence are correlated. More complex tests involve a larger number of processes, thereby increasing the probability of having more processes in common with other tests and consequently showing higher correlations with other tests—a necessary corollary of highly *g*-loaded tests. This would also explain why more complex tasks show greater variance than less complex tasks. Each of the elemental processes that enters into the task performance contributes an independent source of variance, and since variances are additive and since more complex tasks involve more processes, they will show greater variance than less complex tasks. This holds true, of course, only provided that performance on the complex tasks has not become highly automatized by extensive practice.

**3.1.3.** *A Critical Test of the Unitary versus Multiple Processes Hypoth-*

*eses.* A methodology proposed[2] for deciding between the unitary and multiple hypotheses is based on the following argument: If two or more *uncorrelated* variables, e.g., A, B, C, are significantly correlated with another variable, X, then X cannot be unitary, but must contain within it components of variance in common with each of the independent variables A, B, C. This logic can be used to test the unitary hypothesis of *g* by the following steps:

(1) Obtain a good hierarchical *g* by factor analyzing a fair-sized battery of diverse psychometric tests in a sizeable group of subjects.

(2) Obtain RTs, MTs, and measures of intraindividual variability in RT and MT (i.e., SDRT and SDMT) on a fair number of diverse ECTs that tap a number of different information processes.

(3) Do a principal components analysis on all of the measures derived from the ECTs. For the present purpose, principal components analysis has the distinct advantage, over any type of common factor analysis (e.g., principal factors, or principal axes), of yielding component scores (analogous to factor scores) that are perfectly uncorrelated. Factor scores derived from a common factor analysis can be correlated to some extent, even though the factors themselves are perfectly orthogonal (i.e., uncorrelated). (This is because factor scores can never be determined exactly, but can only be estimated and therefore may contain correlated error components.)

(4) For every subject, obtain a *g* factor score (based on the psychometric battery) and principal component scores based on all of the significant components (i.e., those with eigenvalues > 1) derived from all of the ECT variables.

(5) Use the principal component scores as the independent variables in a stepwise multiple regression analysis, with *g* factor scores as the dependent variable.

(6) If more than one (*any* one) of the principal components adds a statistically significant increment to the multiple correlation, the unitary theory of *g* is thereby refuted. The only possible escape from this conclusion would be to claim that the *g* contained "impurities" (i.e., some reliable component of non-*g* variance). This could occur if the *g* were obtained from an insufficiently diverse battery of tests. (For example, the general factor in a battery comprised entirely of various *verbal* tests, such as vocabulary, similarities, analogies, synonyms-antonyms, etc., would not be *g* per se, but an amalgam of *g* and a verbal ability factor.) Of course, if no more than one principal component made a significant contribution to the multiple correlation, it could not prove that the unitary theory of *g* is true, because one could always argue that the "right" ECTs had not been included in the principal component analysis. A mounting number of studies based on more and different ECTs could only increase the likelihood that the unitary theory is true.

The unitary hypothesis was empirically tested using this methodology.[33] From a hierarchical factor analysis of a battery of eleven diverse psychometric tests, *g*-factor scores were obtained in a group of 101 college students. They were also administered eight different ECTs, including simple, choice, and discrimination RT, inspection time, visual scan, memory scan, and speed of retrieving highly over-learned semantic information from long-term memory. Each of the tasks yielded RT and MT and their intraindividual variabilities, SDRT and SDMT, making a total of 32 variables. These were subjected to a principal components analysis, and component scores on the 10 largest components were obtained for every subject. It turned out that four of these components, when entered into a stepwise multiple regression, each contributed significant increments to the prediction of *g*. The simple correlation (Pearson *r*) of the first principal component factor scores with the psychometric *g* factor was .38; the multiple correlation (*R*) based on all four of the principal components that added significant increments to the *R* was .54. Corrected for restriction of range of IQ in this college sample, $R = .66$.

This outcome seems to contradict the hypothesis that *g* has a unitary basis, but the issue is still arguable.[34,35]

But the first principal component of the ECTs, which accounts for by far most of the correlation of the ECTs with *g*, represents mainly a general speed factor, on which the various RTs have the largest loadings. Hence it is important to note that not every (independent) significant component of all of the variance in ECTs is anywhere near being equally related to *g*. The two most *g*-related independent components that emerged from the analysis are RT, which reflects speed of information processing, and SDRT, which reflects intertrial variability or oscillation in processing speed. It is especially noteworthy that the factor—extracted from all of the ECT variables—that mainly represents MT in the various ECTs is quite unrelated to *g* ($r = .03$). Thus MT, when represented by a factor that is completely uncorrelated with RT and SDRT, apparently reflects motor speed and dexterity and has no relation to *g* or information processing. This finding contradicts attempts to explain the correlation between various ECTs and *g* in terms of individual differences in the degree of effort or motivation subjects bring to both types of tasks, since there is no basis for supposing that motivational differences would have an effect on RT while having no effect on MT.

### 3.2. *The Divided Nature of RT*

The same data also throw light on the puzzle, mentioned previously (section 2.3.1), of the ceiling on the correlation between RT and *g*. The largest multiple correlations reported between the RTs based on a number of diverse ECTs and psychometric *g* top out near .65, or about 40% of the *g* variance accounted for by the optimally weighted combination of RTs from several ECTs representing different information processes. Corrections for attenuation and restriction of range seldom raise this correlation much above .70.

Why should there be this apparent ceiling on the size of the RT-*g* correlation? One obvious hypothesis, which has been most clearly spelled out by Detterman,[36] is that *g* arises from the variance contributed by a number of uncorrelated elemental processes. If individual differences in each process were correlated, say, .30 with *g*, thus each accounting for $.30^2 = .09$ of the true variance in *g*, then 11 such independent processes would be needed to account for nearly *all* of the true variance in *g*. This plausible hypothesis seems to have two main problems: (1) it is essentially not testable unless one can account for very nearly all of the true *g* variance by some reasonably small number of uncorrelated process measures or uncorrelated factors derived from a number of ECTs that reflect a large number of independent processes, and it seems doubtful that independent components of variance derived from various ECTs that are devised to reflect all of the different information processes that have been hypothesized to exist would add up to accounting for much more than about 50% of the *g* variance; and (2) probably because of the substantial correlations between process measures, there appears to be a curve of rapidly diminishing returns from including more and more ECTs to a multiple regression equation for predicting *g*, such that the percentage of variance in *g* accounted for by the ECTs asymptotically approaches some value closer to 50% than to 100%. Hence the multiple-process hypothesis of the RT-*g* ceiling needs to be supplemented by another hypothesis to explain the ceiling on the RT-*g* correlation.

A supplementary explanation of the ceiling on the RT-*g* correlation is suggested by these data. When the entire battery of psychometric tests and the ECT variables are factor analyzed together in a hierarchical analysis, there emerges a large second-order *g* factor on which all of the psychometric tests and the RTs and SDRTs of the ECTs are substantially loaded. There is a direct relationship between the complexity of the ECTs and the size of their loadings on this factor. (All MTs have near-zero loadings on this factor.) But here is the most important point: there is also a first-order factor on which only the RT measures, and to a lesser degree MT measures, are substantially loaded, while all of the psychometric tests have near-zero loadings on this factor. It could be called a "noncognitive RT" factor. Also, there is a direct *negative* relationship between the complexity of the ECTs and the sizes of their RT loadings on this factor.

The total common factor variance associated with RTs is divided about half and half between the *g* factor and the noncognitive RT factor. The *g* variance on the RTs and SDRTs of all of the ECTs constitutes, on average, about 40% of the communality (i.e., the common factor variance) of the RTs and SDRTs. But the *g* variance of RT and SDRT averages only about 25% of the *total* variance of RT and SDRT, which is equivalent to an average *g* loading of about .50. This value, if corrected for attenuation and restriction of range, would estimate the true ceiling of the RT-*g* correlation in the general population. It would be close to .70.

The noncognitive components of RT, which are presumably independent of the speed of information processing per se, probably reflect the sensory-motor aspects of RT. (Although the *time* taken up by sensory-motor functioning may constitute only a small fraction of the total RT, especially RT on relatively complex ECTs, it could constitute a considerably larger fraction of the total *variance* in RT.) Hence the ceiling on the RT-*g* correlation is fixed by the proportion of the total RT variance that is attributable to three sources: (1) the noncognitive RT factor, (2) an ECT-specific factor, and (3) measurement error. If these three sources of variance could be eliminated from RTs and SDRTs obtained on a diverse battery of ECTs, their multiple correlation with *g* might well approach unity.

### 3.3. *Information Processing: Top-Down or Bottom-Up?*

This has been another central question. The two main rival hypotheses explain the RT-*g* correlation quite differently.

(1) The *top-down* hypothesis holds that the *higher-level* mental processes, which are obviously required for the kinds of complex problem solving seen in the most highly *g*-loaded psychometric tests, govern an individual's RT on even relatively simple tasks via strategies, the so-called executive or metaprocesses, and "attentional resources." Therefore, the causal locus of individual differences in RT, as in *g*, is in the higher-level mental processes, which critically affect the lower-level processes manifested in such variables as choice RT and IT (inspection time).

(2) The *bottom-up* hypothesis does not deny the existence of high-level processes, of course, but holds that there are stable individual differences in relatively simple but pervasive neural processes, such as nerve conduction velocity and synaptic delay, which affect the speed and efficiency of the transmission of information in the CNS, and that these properties are involved to some degree at *all* levels of information processing, from that of the simplest tasks such as choice RT and IT to that of the complex items of conventional IQ tests. Hence there is a correlation between individual differences in, for example, discrimination RT and the *g* of complex psychometric tests.

**3.3.1.** *Nerve Conduction Velocity (NCV) in the CNS and* g. Investigations by Reed[37] of the heritability of individual differences in NCV in mice led to his hypothesizing NCV as a physiological basis for the heritability of *g*.[38,39] A critical test of this hypothesis produced results consistent with the bottom-up hypothesis.[40] Short-latency visually evoked potentials (VEPs N70 and P100) in response to pattern-reversal stimulation and recorded over the primary visual cortex showed highly reliable individual differences in 147 college males. The latencies of the earliest clearly-defined neural impulses to be transmitted from the retina through the visual tract to the visual cortex are very short—70 to 100 msec. In order to obtain approximate measures of individual differences in nerve conduction velocity, i.e., distance/time, it was necessary to take into account individual differences in head size. The estimate of an individual's NCV was obtained by dividing his head length by the mean latency of the VEP. The resulting approximate measures of NCV (here labeled V:P70 and V:P100) were found to be significantly correlated with IQ scores on Raven's Advanced Progressive Matrices, a highly *g*-loaded nonspeeded, nonverbal test of complex reasoning ability. The correlation for V:N70 was +.18 ($p$ = .025) and for V:P100, +.26 ($p$ = .002). Correction for the restriction of range of IQ in the college sample raises these correlations to +.27 and +.37, respectively. FIGURE 1 shows the mean IQ of these college students at each quintile of the V:P100.

The theoretical significance of this finding is based on the fact that the latencies of neural impulses through the visual tract and recorded on the visual cortex are very much shorter than the total amount of time needed for the neural impulses to reach the higher cortical centers involved in solving Raven Matrices problems, and therefore the VEP latencies could not be affected by any top-

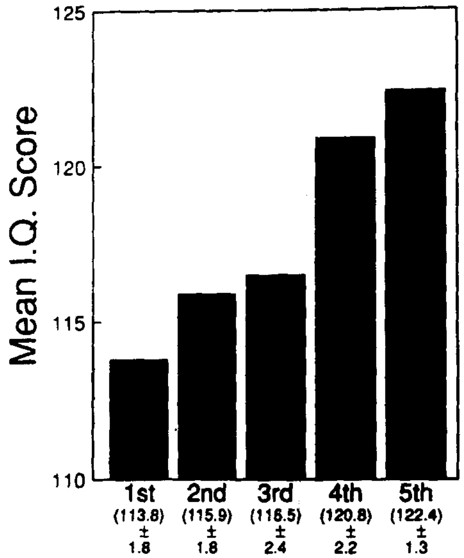

**FIGURE 1.** Distribution of mean IQ scores (±SE) in V:P100 quintiles. The distribution of V:P100 values (i.e., the NCV based on the P100 latency) of the 147 students, from the lowest NCV (1.75 m/sec) to the highest (2.22 m/sec) was divided into quintiles. Quintile 1 contains the 20% of students with the lowest V:P100 values, quintile 2 contains the 20% of students with V:P100 values between the 20th and 40th percentiles, etc. 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 linearity.

down processes. The explanation for the observed correlation between NCV in the visual tract and *g* rests 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 themselves very similar, and thus individual differences in visual tract NCVs and cortical NCVs are correlated (Reed and Jensen, submitted for publication). Because information is transferred from one cortical region to another via axons at some velocity and across synapses with some delay, the mean cortical 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, seem a most plausible component of *g*.

This hypothesis does not in the least contradict the obvious necessity of hypothesizing specific neural structures and their complex functional organization or patterning to explain the facts about information processing. At present we have virtually no knowledge of the extent to which these design features of cortical functioning contribute to individual differences in cognitive abilities. But we do have evidence now which suggests that brain NCV alone may account for a substantial part, perhaps as much as 25%, of the *g* variance in the general population.

**3.3.2.** *Peripheral Nerve Conduction Velocity and* g. It is still quite uncertain whether NCV in *peripheral* afferent or efferent nerves is correlated with *g*. A positive correlation was first hypothesized by Reed,[38] on the supposition that NCV in peripheral nerves would have some relationship to cortical NCV. Three independent studies that were expressly designed to test this hypothesis have shown seemingly contradictory results.

First, Vernon and Mori[18] measured NCV in the median nerve of the forearm in 85 college students and found a correlation of +.42 between NCV and *g*, and a correlation of −.28 between NCV and RT. (RT and *g* were correlated −.44.) Also, the higher the *g* loadings of the ten subtests in a multi-aptitude battery, the higher were the subtests' correlation with NCV. It could hardly be more beautiful!

But then another study[25] found no significant correlation between NCV (measured via sensory nerve action potentials in the median nerve between the middle finger and the wrist) and IQ (Raven Advanced Progressive Matrices) in a quite heterogeneous group of 44 adults. However, significant correlations were found between the intertrial variabilities (i.e., within-subject SDs) of the sensory nerve action potential latencies and IQ. When the variabilities of the median nerve action potential latencies measured in the right and left hands were entered into a multiple regression to predict IQ, the overall shrunken multiple *R* is about .50. The authors[25] caution that their correlations "must be viewed as no more than exploratory values that require replication" (p. 10).

The most recent study,[41] based on 200 male college students between 18 and 25 years of age, found completely nonsignificant and near-zero correlations between NCV in the median nerve of the forearm and the following variables: (1) IQ (Raven Advanced Progressive Matrices), (2) RT on three different ECTs of increasing complexity, and (3) brain NCV (i.e., visual pathway, from retina to occipital lobes).

The discrepant results of these three studies have no obvious explanation, although it should be noted that they differ in their methods for measuring NCV in the median nerve. Just how these procedural differences could affect the results is still mysterious. Replications are obviously essential, and Vernon and Mori (personal communication) are presently repeating their study.

### 3.4. *The Relation between Speed of Information Processing and the Capacity of Working Memory*

If there is a principal focus of *g* variance in the processing system, most cognitive theorists would probably point to *working memory* (WM), a hypothetical construct regarded as the central information processing unit. A component of short-term memory (STM), WM comprises the functions of focusing attention, conscious rehearsal, and transformation and mental manipulation of information received from external sources or retrieved from long-term memory (LTM). WM is the *active* aspect of STM. Some theorists do not distinguish between the *passive* and *active* aspects of STM, but it is theoretically useful to do so. STM comprises both *primary* memory and *working* memory. Primary memory is the *passive* holding station for recent input. The WM plays the *active* role, manipulating or transforming the input.

Both primary memory and WM are called short-term memory (STM) systems because the neurally encoded traces of information in them undergo rapid loss and become inaccessible within a matter of seconds, unless they are continuously rehearsed until consolidated in LTM. Transferring information from STM into LTM is one of the functions of WM. The storage capacity of LTM is practically unlimited. But a crucial characteristic of STM is its very limited capacity. This limitation means that WM can deal with only a certain small amount of information per unit of time, which some theorists express as number of bits per second. (A *bit* is the binary logarithm of the number of alternative choices in a decision process, or the amount of information needed to reduce uncertainty by one-half). An excessive rate of information input overloads the capacity of WM, causing a "breakdown" of processing and a loss of information. Who has not looked up a phone number, and then, just before dialing it, is asked a question that demands a quick answer?

Empirically, there is an intimate relationship between the *capacity* limitation of WM and the *speed of processing* as measured by RT tasks. In a so-called *dual task* paradigm, in which the subject must attend to two distinct tasks either simultaneously or in quick succession, RT appears to be a sensitive indicator of capacity. Significantly, the RT in such dual tasks turns out to be more highly correlated with *g* than the RTs on either of the single tasks. For example, if a person is presented a set of several digits to retain for later recall, then we immediately require the person to perform a choice RT task, and then cue the person to recall the digits, we find that the person's RT is significantly longer, on average, than when RT is measured by the same choice RT task when it is unaccompanied by the digit span task. It is as if RT is at least partly a function of the available capacity of WM. The more fully WM is occupied, the slower is the reaction to an external stimulus.

*Capacity* seems a necessary concept for understanding *g*, because persons do not show the same rank order in ability on tasks that place greater or lesser amounts of strain on WM. If there were just a single cause of individual differences, say, speed of processing, then we should expect a perfect (disattenuated) correlation between single-task RT and dual task RT, for example. Yet there is not a perfect correlation. Another example is the comparison of forward digit span (FDS) and backward digit span (BDS): they are not perfectly correlated. BDS is subjectively harder than FDS, fewer digits can be recalled in the BDS than in the FDS task, and apparently the mental operation of reversing the digit series in BDS takes up more of the capacity of WM than is taken up by FDS. Interestingly, BDS has almost double the correlation with IQ as FDS.[42]

Most cognitive theorists believe that at least two distinct fundamental variables are needed to explain *g*: (1) the *speed* of information processing, and (2) the *capacity* of STM (including WM). But how can we conceptualize the connection between capacity and processing speed? Psychologists in Erlangen, Germany, have given this problem the most attention so far.[43] They argue essentially that the capacity (C) of STM (expressed in bits of information) is the product of the speed (S) of information flow (or processing speed, in bits per second) and the duration time (D) of information in STM absent rehearsal. That is, C bits = S bits/sec x D sec. Assuming that the parameters S and D are to some degree independent sources of variance, this formulation is consistent with the previously described phenomena related to dual tasks and to forward and backward digit span.

Experiments by the Erlangen psychologists have empirically obtained estimates of these parameters in average adults, approximately, of S = 15 bits/sec, D = 5 to 6 sec, and C = 80 bits. Accordingly, for this theory, a measure of the capacity of WM should be more highly correlated with *g* than is a measure of processing speed alone. Lehrl and Fischer[43] report correlations of .67 and .88 between their experimental measure of STM capacity (as $C = S \times D$) and scores on a vocabulary test in two samples of adults, with *N*s of 672 and 66, respectively. (Vocabulary is loaded on a verbal factor but is also typically the most highly *g*-loaded test in psychometric batteries.) Thus individual differences in measurements derived from exceedingly simple tests that are virtually devoid of intellectual content but which provide estimates of processing speed and the duration of information in STM are correlated to a remarkable degree with individual differences in an especially complex and highly *g*-loaded cognitive ability such as vocabulary.

## 4. *Structural and Design Features of the Brain*

After all that has been said about *speed* of information processing and its neural correlates, we must recognize that there are probably other aspects of the physical basis of *g* that are possibly of equal or even greater importance. The main problem at present, however, is our almost total ignorance of the extent of individual differences in the structural features of the brain and the degree to which they are related to *g*.

Theory and research on the relation of anatomical, physiological, and neural architectonic aspects of the brain to mental abilities and other types of behavior have focused mainly on particular classes of performance (e.g., language, memory, hemispheric functions) rather than on individual differences, at least among neurologically intact persons whose cognitive abilities are within the normal range. Although this nomothetic aspect of brain research, focused primarily on the localization and structural substrate of psychological functions, is extremely important in its own right, it affords few clues at present about the nature of *g*. With the rapid advances in this branch of neuroscience, however, it should soon become a source of potentially (though not easily) testable hypotheses concerning those design aspects of the brain that might explain some part of *g* not accounted for by general speed of information processing. I suspect, however, that individual differences in localized neural structures will much more likely account for the so-called *group factors*, such as verbal, spatial, and musical abilities, than for *g* per se.

Many specialized abilities displayed by humans are so complex in terms of the amount of information processing involved that a linear model of information transmission in which NCV is the main source of variance would be entirely inadequate to explain the rapidity of performance times for highly complex tasks. For example, Shakuntala Devi, the famous calculating prodigy, is able mentally to extract cube roots (or many other roots) of enormous numbers in the hundreds of millions in just a few seconds, yet she is not especially exceptional either in *g* or in speed of information processing on a variety of ECTs.[44] Hence it is necessary to hypothesize certain structural design features of the neural substrate, such as parallel processing, hierarchical organization of neurones, automatization of selecting and executing complex algorithms, and the like. That these features of brain organization may be an important source of individual differences, at least for specialized abilities if not for *g*, is suggested by the astounding performances of prodigies in mathematics, music, and chess.[45]

Willerman and Raz[46] have hypothesized several distinct features of the brain that might contribute to *g*. Most basic, perhaps, is the number of cortical cell analyzers, including their organization and degree of interconnectedness as reflected by the size or number of association and commissural nerve fibers in the white matter underlying the cerebral cortex. It is noted that, compared with other species, humans have a higher proportion of cortical white matter in relation to gray matter. The white matter consists of the myelinated fibers of the cortical neurons, which are called association fibers, as they interconnect neurons within each hemisphere. The myelinated commissural fibers interconnect neurons across the two hemispheres. Willerman and Raz hypothesize as follows: "The association and commissural fibers would seem especially relevant to intelligence because they probably represent important neurological underpinnings for educing relations and correlates, or more simply put, making connections between disparate cognitive elements and applying them to new problems" (p. 9). But there have been no systematic studies of individual differences in the size or number of association fibers, or of their relation to *g*..

Spearman originally discovered that pitch discrimination is correlated with *g*, and his finding has been confirmed by subsequent studies.[47] This fact would seem hard to explain in terms of individual differences in speed of information processing or NCV per se. Noting this problem, Willerman & Raz[46] propose the following hypothesis: "Neurons varying in central frequencies and bandwidths can be organized to produce many features of intelligent functioning as indexed by our discrimination experiments. In the auditory domain, one neuron might be especially sensitive to hair cells that have a modal firing frequency of 800 to 810 Hz, another sensitive to frequencies of 805 to 815 Hz. In addition, their conjoint activation could trigger a third neuron, . . . referred to as a 'grandmother' cell. An activated grandmother cell indicates that the specific frequency of the sensory signal must have been between 805 and 810 Hz. By adding increasing layers to the hierarchy of neurons, for example, great grandmother cells that fire only when two grandmother cells are simultaneously activated, the frequency of the original signal can be retrieved with increasingly greater precision. Therefore, it seems reasonable to propose the theory that intelligent people have a greater number of cortical elements arranged in some hierarchical order which permits finer analysis of signals" (p. 9). Of course, the fact that pitch discrimination is *g*-loaded suggests that there must be other such hierarchical neural structures, probably of common epigenetic origin and hence with correlated individual differences, that play a part in many other forms of information processing besides auditory discrimination.

Finally, Willerman and Raz[46] consider biochemical factors as possibly contributing to *g*. They exclude specific neurotransmitters, as these vary widely and seem to be keyed to localized regions and specialized functions, while *g* implies a much broader causal factor. But they hypothesize that "certain ions (e.g., $K^+$ and $Ca^{++}$) are ubiquitous in all forms of chemical neurotransmission and variations in their concentrations and transport might be candidates for an endogenous physiological process underlying the general factor in intelligence" (p. 10).

## 5. *Conclusion*

Psychometrically, the existence of *g* is an established fact. Its causal nature, however, is still far from being fully understood, although recent research aimed toward this goal has brought forth fairly impressive empirical evidence that *g* is more closely related to a biological state of affairs than probably any other mental ability factors independent of *g*. Considerable evidence also supports the theory that a large part of *g* is attributable to individual differences in the speed of information processing, with its physiological counterpart, neural conduction velocity in the brain, which is also found to be correlated with *g*. Evidence now favors, but has not proved decisively, the hypothesis that *g* is not unitary in the sense of being attributable to some single property of the brain, but comprises independent sources of variance that must be attributable to a probably quite limited number of distinct brain processes. Recent advances in the neurosciences afford several plausible hypotheses, as yet untested, of structural and biochemical properties of the brain that could possibly contribute to *g*.

Hence the prospect of achieving an empirically valid theory of the causal nature of *g* now looks remarkably promising. And if there is still any doubt that important strides have been made in this direction since Spearman's day, it should prove instructive to compare the present picture with the rather faint-hearted and pessimistic view of this endeavor envisaged by Spearman.[48] He wrote as follows: "We have introduced no hypothesis as to the essential nature of what is measured by *g*. . . . But for scientific ends, there is much advantage in doing so. For the purpose of building up an intelligible whole, and also for that of inspiring further investigation, there is urgent need of framing—however tentatively and provisionally—some or other explanatory hypothesis" (p. 414). "And even should the worst arrive and the required physiological explanation remain to the end undiscoverable, the mental facts will none the less remain facts still. If they are such as to be best explained by the concept of an underlying energy, then this concept will have to undergo that which after all is only what has long been demanded by many of the best psychologists—it will have to be regarded as purely mental" (p. 408).

It would be most surprising indeed if Spearman would have entertained the thought of such an unsatisfactory outcome for the theory of *g* if he had foreseen the findings of the research summarized in this paper. In the next two or three decades, as progress in neuroscience continues apace, the kind of *g* theory that Spearman hopefully envisaged might well be empirically substantiated—in his words, "whereby physiology will achieve the greatest of all its triumphs" (p. 407).

## 6. *Summary*

Individual differences with respect to diverse tests of mental abilities that range in complexity from simple reaction time to abstract reasoning are all positively correlated in the population. The total covariance among all such tests can be analyzed into a number of uncorrelated components of variance, or factors, that, in terms of their generality, are hierarchical, with the most general factor, or *g*, at the apex. This *g* factor is common to every type of cognitive performance, whatever other ability factors may be involved (e.g., verbal, spatial, numerical, musical, etc.), and is the crucial factor in most tests' practical validity. Its correlations with various tests' heritability, inbreeding depression, heterosis, average evoked potentials, brain metabolism, and many other physical correlates indicate that as a product of evolution it is profoundly enmeshed with many organismic variables. A theory based on empirical evidence links *g* to neural processes involved in the speed and efficiency of information processing.

## REFERENCES

1. Spearman, C. 1904. "General Intelligence," objectively determined and measured. Am. J. Psychol. **15:** 201–293.
2. Jensen, A. R. 1987. Psychometric *g* as a focus of concerted research effort. Intelligence **11:** 193–198.
3. Jensen, A. R. Phlogiston, animal magnetism, and intelligence. *In* Theories of Intelligence. D. K. Detterman, Ed. Ablex Publishing Corp. Norwood, NJ. In press.
4. Thorndike, R. L. 1987. Stability of factor loadings. Pers. Indiv. Differ. **8:** 585–586.
5. Jensen, A. R. 1988. The relationship between learning and intelligence. Learn. Indiv. Differ. **1:** 37–62.
6. Gottfredson, L., Ed. 1986. The *g* factor in employment. J. Vocat. Behav. **29**(3).
7. Jensen, A. R. 1991. Psychometric *g* and achievement. *In* Current Views on Testing: A Policy Perspective. B. R. Gifford, Ed. Kluwer Academic Publishers. Boston, MA.
8. Thorndike, R. L. 1985. The central role of general ability in prediction. Multivariate Behav. Res. **20:** 241–254.
9. Thorndike, R. L. 1986. The role of general ability in prediction. J. Vocat. Behav. **29:** 332–339.
10. Jensen, A. R. 1986. *g*: Artifact or reality? J. Vocat. Behav. **29:** 301–331.
11. Jensen, A. R. 1987. The *g* beyond factor analysis. *In* The Influence of Cognitive Psychology on Testing and Measurement. J. C. Conoley, J. A. Glover & R. R. Ronning, Eds. Lawrence Erlbaum. Hillsdale, NJ.
12. Jensen, A. R. 1983. Effects of inbreeding on mental ability factors. Pers. Indiv. Differ. **4:** 71–87.
13. Schafer, E. W. P. 1985. Neural adaptability: A biological determinant of *g* factor intelligence. Behav. Brain Sci. **8:** 240–241.
14. Jensen, A. R. & S. N. Sinha. 1993. Physical correlates of human intelligence. *In* Biological Approaches to the Study of Human Intelligence. P. A. Vernon, Ed. Ablex Publishing Corp. Norwood, NJ.
15. Jensen, A. R. 1980. Uses of sibling data in educational and psychological research. Am. Educ. Res. J. **17:** 153–170.

16. Cohn, S. J., C. M. G. Cohn & A. R. Jensen. 1988. Myopia and intelligence: A pleiotropic relationship? Hum. Genet. **80:** 53–58.
17. Willerman, L., R. Schultz, J. N. Rutledge & E. D. Bigler. 1991. *In vivo* brain size and intelligence. Intelligence **15:** 223–228.
18. Vernon, P. A. & M. Mori. 1989. Intelligence, reaction times, and nerve conduction velocity [Abstract]. Behav. Genet. **19:** 799.
19. Jensen, A. R., S. J. Cohn & C. M. G. Cohn. 1989. Speed of information processing in academically gifted youths and their siblings. Pers. Indiv. Differ. **10:** 29–34.
20. Jensen, A. R., G. E. Larson & S. Paul. 1988. Psychometric *g* and mental processing speed on a semantic verification test. Pers. Indiv. Differ. **9:** 243–255.
21. Jensen, A. R. & T. E. Reed. 1990. Simple reaction time as a suppressor variable in the chronometric study of intelligence. Intelligence **14:** 375–388.
22. Kranzler, J. H. & A. R. Jensen. 1989. Inspection time and intelligence: A meta-analysis. Intelligence **13:** 329–347.
23. Vernon, P. A., Ed. 1987. Speed of Information-Processing and Intelligence. Ablex Publishing Corp. Norwood, NJ.
24. Vernon, P. A. & M. Mori, M. 1990. Physiological approaches to the assessment of intelligence. *In* Handbook of Psychological and Educational Assessment of Children: Intelligence and Achievement. C. R. Reynolds & A. W. Kamphaus, Eds. Guilford Press. New York.
25. Barrett, P. T., I. Daum & H. J. Eysenck. 1990. Sensory nerve conduction and intelligence: a methodological study. J. Psychophysiol. **4:** 1–13.
26. Mathews, G. & L. Dorn. 1989. IQ and choice reaction time: An information processing analysis. Intelligence **13:** 299–317.
27. Kutas, M., G. McCarthy & E. Donchin. 1977. Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. Science **197:** 792–795.
28. Vernon, P. A. 1989. The heritability of measures of speed of information-processing. Pers. Indiv. Differ. **10:** 573–576.
29. McGue, M., T. J. Bouchard, Jr., D. T. Lykken & D. Feuer. 1984. Information processing abilities in twins reared apart. Intelligence **8:** 239–258.
30. Ho, H.-Z., L. Baker & S. N. Decker. 1988. Covariation between intelligence and speed-of-cognitive processing: Genetic and environmental influences. Behav. Genet. **18:** 247–261.
31. Jensen, A. R. 1987. Process differences and individual differences in some cognitive tasks. Intelligence **11:** 107–136.
32. Willerman, L. & J. M. Bailey. 1987. A note on Thomson's sampling theory for correlations among mental tests. Pers. Indiv. Differ. **8:** 943–944.
33. Kranzler, J. H. & A. R. Jensen. 1991. The nature of psychometric *g*: Unitary process or a number of independent processes? Intelligence. **15:** 379–422.
34. Carroll, J. B. 1991. No demonstration that *g* is not unitary, but there is more to the story: Comment on Kranzler and Jensen. Intelligence. **15:** 423–436.
35. Kranzler, J. H. & A. R. Jensen. 1991. Unitary *g*: Unquestioned postulate or empirical fact? Intelligence. **15:** 437–448.
36. Detterman, D. K. 1987. What does reaction time tell us about intelligence? *In* Speed of information-processing and intelligence. P. A. Vernon, Ed. Ablex Publishing Corp. Norwood, NJ.
37. Reed, T. E. 1988. Narrow-sense heritability estimates for nerve conduction velocity and residual latency in mice. Behav. Genet. **18:** 595–603.
38. Reed, T. E. 1984. Mechanism for heritability of intelligence. Nature **311:** 417.
39. Reed, T. E. 1988. A neurophysiological basis for the heritability of intelligence. *In* Intelligence and Evolutionary Biology. H. J. Jerison & I. Jerison, Eds. Springer-Verlag. Berlin.

40. Reed, T. E. & A. R. Jensen. 1993. Conduction velocity in a brain nerve pathway of normal adults correlates with intelligence level. Intelligence. **16:** 259–278.
41. Reed, T. E. & A. R. Jensen. 1991. Arm nerve conduction velocity (NCV), brain NCV, reaction time, and intelligence. Intelligence **15:** 33–47.
42. Jensen, A. R. & R. A. Figueroa. 1975. Forward and backward digit-span interaction with race and IQ: Predictions from Jensen's theory. J. Educ. Psychol. **67:** 882–893.
43. Lehrl, S. & B. Fischer. 1988. The basic parameters of human information processing: Their role in the determination of intelligence. Pers. Indiv. Differ. **9:** 883–896.
44. Jensen, A. R. 1990. Speed of information processing in a calculating prodigy. Intelligence **14:** 259–274.
45. Obler, L. K. & D. Fein, Eds. 1988. The exceptional Brain: Neuropsychology of Talent and Special Abilities. Guilford Press. New York.
46. Willerman, L. & N. Raz. 1987. Implications of research on auditory recognition masking for a theory of intelligence. Paper presented at the 3rd meeting of the International Society for the Study of Individual Differences, in Toronto, Canada, June 18–22.
47. Lynn, R., G. Wilson & A. Gault. 1989. Simple musical tests as measures of Spearman's *g*. Pers. Indiv. Differ. **10:** 25–28.
48. Spearman, C. 1927. The Abilities of Man: Their Nature and Measurement. AMS Press. New York.