ADVANCES IN COGNITION AND EDUCATIONAL PRACTICE

Volume 4 • 1997

REFLECTIONS ON THE CONCEPT OF INTELLIGENCE

ADVANCES IN COGNITION AND EDUCATIONAL PRACTICE

REFLECTIONS ON THE CONCEPT OF INTELLIGENCE

Series Editor: JERRY S. CARLSON
University of California
Riverside

Volume Editors: JOHANNES KINGMA
Department of Traumatology
University Hospital Groningen
The Netherlands

WELKO TOMIC
Faculty of Social Sciences
The Open University
The Netherlands

VOLUME 4 • 1997

JAI PRESS INC.

Greenwich, Connecticut *London, England*

55 Old Post Road No. 2
Greenwich, Connecticut 0680

JAI PRESS LTD.
38 Tavistock Street
Covent Garden
London WC2E 7PB
England

ISBN: 0-7623-0105-8

Manufactured in the United States of America

CONTENTS

LIST OF CONTRIBUTORS

C. Davison Ankney	Department of Zoology University of Western Ontario
María Dolores Calero	Hospital Real University of Granada
John Colombo	Department of Human Development University of Kansas
Roberto Colom	Faculty of Psychology University of Madrid
H.J. Eysenck	Institute of Psychiatry University of London
Rocío Fernández-Ballesteros	Faculty of Psychology University of Madrid
Arthur R. Jensen	School of Education University of California
Manuel de Juan-Espinosa	Faculty of Psychology University of Madrid
Johannes Kingma	Department of Traumatology University Hospital Groningen The Netherlands
Maria Levander	Department of Neurology Korolinska Hospital Sweden
Charles Locurto	Department of Psychology College of the Holy Cross

Gerald Matthews

Department of Psychology
University of Dundee
United Kingdom

D. Wayne Mitchell

Southwest Missouri State University

A. C. Neubauer

Department of Psychology
University of Graz
Austria

J. Philippe Rushton

Department of Pychology
University of Western Ontario

Welko Tomic

Faculty of Social Sciences
The Open University, Heerlen
The Netherlands

PREFACE

We can hardly mention a human characteristic that is more inextricably bound up with human existence than intelligence. Since time immemorial, man has observed his own species in everyday situations and concluded that not every one of his fellow humans is endowed with the same degree of intelligence. As far back as the ninth or eighth century B.C., Homer noted that intelligence is not meted out in equal proportions. He wrote:

> So it is that the gods do not give all men gifts of grace—neither good looks nor intelligence nor eloquence (Homer VIII, p. 167).

For centuries man has attempted to get to the bottom of the phenomenon of intelligence. Nevertheless, it has not become any easier to clarify what we mean by intelligence or to define the concept of intelligence in a way that is acceptable to all experts. Van Peet (1992), for example, traced 73 different definitions of the concept in the literature.

Despite the different definitions there are a number of recurrent issues in the literature on intelligence which emerged of the beginning of this century and have continued right through to contemporary research. This volume focuses on two such issues: general intellectual capacity, the *g* factor (and elaborations on it), and how to influence the development of intelligence.

With regard to general intellectual capacity, Spearman (1923, 1927) devised this term at the beginning of this century to describe an ability he felt

underpinned all intellectual skills. The founder of a major psychometric theory, Spearman showed that individuals who did well on one mental ability test tended to do well on other tests as well. The reverse also seemed to be true. Spearman used the technique of factor analysis to examine individual differences in test scores. He was convinced that such analysis revealed the underlying sources of individual differences. The method he developed suggested to him that there were two factors behind such differences. The most important of the two he called "*g*", the general factor, which explains achievement on all tasks requiring intelligence. The *g* factor is responsible for the intercorrelation of different intelligence tests. It is a common factor linking all intellectual performance.

Another issue within the context of research into the *g* factor is that intelligence has been viewed historically as a more or less fixed trait. It was thought to be a general ability which was innate, fairly constant over time and measurable by tests administered to groups or to individuals. Researchers who take this view perceive intelligence as a trait people are born with. Certainly there are very few researchers who would deny that development also plays a role; in their view, development allows this genetic endowment to express itself. One notable method to assess the heritability of intelligence is to study identical twins who have been raised apart. There is some agreement that estimates of heritability can differ among ethnic and racial groups. However, it should be noted that such estimates are based on test scores and only cover that part of intelligence measured by those tests.

Differential research into the *g* factor focused on differences between blacks and whites (Jensen, chapter 1), brain size and race (Rushton & Ankney, chapter 2). However, recently some researchers have pointed out the need to investigate the biological underpinnings of *g*, for example, by means of reaction time or physiological measures (Eysenck, chapter 5). Some researchers want to broaden the scope of the *g* factor. For example, in contrast to the idea that *g* may be employed to explain the continuity of intellectual functioning from infancy, research into infant cognition may well elucidate the cognitive processes that comprise *g* in the adult (Mitchel & Colombo, chapter 4). Research into the biological underpinnings of *g* encompass studies on human and animal similarity in *g* (Locurto, chapter 3) as well as reaction times in elementary cognitive tasks (Neubauer, chapter 6), and cognitive assessment procedures for traumatic head injuries and epilepsy (Levander, chapter 8). Other studies have broadened the scope of the *g* factor by observing it as a sub factor of personality as well as by integrating *g* into a theory of multiple intelligence (Matthews, chapter 7).

The second recurrent issue in intelligence research, that is, whether intelligence scores can be increased or whether the development of intelligence can be accelerated, is covered in two chapters in this volume (Fernández-Ballesteros, de Juan-Espinosa, Colon & Calero, chapter 9, and Kingma &

Tomic, chapter 10). At the beginning of this century Binet, the founder of the intelligence test, was already convinced that a child's intelligence level could be raised by teaching it how to learn (Resing, in press). In Europe Binet's ideas continued to influence later scholars (e.g., Selz, 1935, Kohnstamm, 1942), who distinguished between intelligence *in actu* and *in potentia.*

In the early fifties (Braine, 1959) and sixties, many researchers in and outside the United States focused on training children in typical Piagetian concepts. The learning controversy was invoked by Bruner's (1966) debate with Inhelder (a Genevan colleague of Piaget) focusing on the idea of "readiness to learn". After publication of this debate, an enormous number of training experiments were conducted in the developmental psychology literature in the late sixties and early seventies (Brainerd, 1978a, 1978b).

Although some short-term training may have invoked durable effects (for some weeks), the question as to whether intelligence had been raised to a higher level was still unanswered since no golden (stringent) standard had been used to evaluate the long-term training effect.

During the seventies and eighties, European psychologists (Van Parreren, 1988) and scholars in the United States (Wertsch, 1985) emphasized the importance of the child's social environment in the development of intelligence. They agreed with Vygotsky, who suggests that the development of intelligence may be influenced largely by the interaction with others. Feuerstein, Rand, Hoffman, and Miller (1980), for instance, suggest that the key to the development of intelligence is the mediated learning experience. The learning potential and the context in which learning takes place are considered important factors in the development of intelligence (Fernández et al., chapter 9). However, other still unknown factors may also influence a person's intelligence level. Flynn (1984) has shown that intelligence test scores have been rising steadily. He concludes that between 1932 and 1978, the average IQ increased about 14 IQ-points. There is no satisfactory explanation for this phenomenon as yet.

We are pleased to include contributions by scholars from Canada, Europe and the United States, representing diverse viewpoints in the field of research into the *g* factor and into the possibility of raising a person's level of intelligence. The authors summarize their main research topic, enabling the reader to study their work in depth. Chapters 1 through 8 concern *g* factor-related research, chapters 9 and 10 deal with research into influencing the level of intelligence.

Jensen (chapter 1) provides an in-depth summary of his research into differences between black and white performance on psychometric mental ability tests. Samples of the black and white populations from the United States show the same *g*, that is, the general factor common to all measures of cognitive ability. Jensen discusses B-W differences in terms of the *g* factor. In addition he also shows other relationships between *g* and some cultural variables.

Rushton and Ankney (chapter 2) provide a comprehensive review of the research into race and sex differences in brain size and cognitive ability. They show that different techniques have produced convergent evidence that brain size correlates with cognitive ability and varies by race and sex. An explanation for *g* factor differences is discussed in terms of different theoretical points of view.

Locurto (chapter 3) also discusses the general factor in human intelligence. He broadens the scope to the study of animal intelligence. However, to date no general factor has been reported in animal intelligence.

Mitchell and Colombo (chapter 4) discuss the relationship between the *g* factor and infant intelligence. They argue that research into infant information processing may well elucidate the cognitive processes that comprise *g* in the adult.

Eysenck (chapter 5) provides an historical overview of the founders of the scientific study of intelligence, Binet and Galton. He points out differences between the two historical mainstreams of intelligence research. Although Binet's views and methods prevailed, Eysenck shows that in recent years there has been a return to Galton's approach, as demonstrated in new research into the genetic and biological mechanisms underlying *g*. This new perspective emphasizes the importance of physiological measures (e.g., the mental speed approach) when assessing the *g* factor.

Neubauers's review of the mental speed approach (chapter 6) complements Eysenck's new perspective on intelligence, that is, that the speed of information processing is physiologically based. He reviews different types of models and concludes that elementary cognitive tasks may enable a more culture-fair assessment of some kind of "biological intelligence."

Matthews (chapter 7) focuses in his chapter on the cognitive correlates of intelligence and personality. He shows that the *g* factor is more strongly related to certain cognitive functions than others. A similar picture emerges for the cognitive correlates of personality traits. Whereas *g* is almost always associated with processing efficiency, traits are associated with the enhancement of some processes related to adaptive trade off. The conclusion is that intellectual and personality traits are linked to adaptive decisions.

Levander (chapter 8) surveys her research into theoretical and methodological issues concerning adult intelligence within the context of neuropsychology. She describes cognitive assessment procedures for cognitive impairments related to head injuries and epilepsy and how to evaluate cognitive rehabilitation.

Fernández-Ballesteros, de Juan-Espinosa, Colon and Calero (chapter 9) give a comprehensive outline of their research into psychometric and experimental procedures studying individual differences in intelligence and learning. They discuss different levels of general adjustments in the personal context to specific task performance levels, and illustrate the mediating aspect of the learning context as well as the improvement of the subject's intellectual performance. Learning potential measures are shown to be the best predictor for intellectual performance in long-term intervention programs.

Kingma and Tomic (chapter 10) discuss the effects of intervention programs on accelerating the development of intelligence within the context of Piaget's criteria for the assessment of durable training effects. A durable change in the development of intelligence can only be obtained if the context (in which the newly acquired "skills" are exercised) has also been changed, more or less permanently.

Johannes Kingma
Welko Tomic
Volume Editors

REFERENCES

Braine, M.D.S. (1959). The ontogeny of certain logical operations: Piaget's formulation examined by nonverbal methods. *Psychological Monographs, 73* (5), Whole No, 475.

Brainerd, C.J. (1978a). Learning research Piagetian theory. In L.S. Siegel & C.J. Brainerd (Eds.), *Alternatives to Piaget: Critical essays on the theory*. New York: Academic Press.

Brainerd, C.J. (1978b). *Piaget's theory of intelligence*. Englewood Cliffs, NJ: Prentice-Hall.

Feuerstein, R., Rand, Y., Hoffman, M.B., & Miller, R. (1980). *Instrumental enrichment: An intervention program for cognitive modifiability*. Baltimore: University Park Press.

Flynn, J.R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, *95*, 29-51.

Kohnstamm, P. (1942). Over het probleem van "psychische metingen" in 't algemeen en van "intelligentie metingen" in 't bijzonder. [On the problem of "mental measurements" in general and of "intelligence measurements" in particular]. *Tijdschrift voor Philosofie*, *3*, 223-344. [Journal of Philosophy].

Peet, A.A.J. van, (1992). *De potentieeltheorie van intelligentie*. [The potential theory of intelligence]. Proefschrift. [Ph.D. thesis]. Amsterdam: Universiteit van Amsterdam.

Resing, W.C.M. (In press). Intelligence, learning and learning potential. In W. Tomic, H.T. van der Molen, & W.K.B. Hofstee, (Eds.), *Intelligence and social competence*. Lisse: Swets & Zeitlinger.

Selz, O. (1935). Versuche zur Erhebung des Intelligenzniveaus. Ein Beitrag zur Theorie der Intelligenz und ihrer erziehlichen Beeinflüssung. [Trials in increasing the level of intelligence. A contribution to the theory of intelligence and its educational influence]. *Zeitschrift für Psychologie, 134*, 236-301. [Journal of Psychology].

Spearman, C.E. (1923). *The nature of "intelligence" and the principles of cognition*. London: Macmillan.

Spearman, C.E. (1927). *The abilities of man: Their nature and measurement*. New York: Macmillan.

Van Parreren, C.F. (1988). De handelingspsychologische studie van de cognitieve ontwikkeling. [Action theoretical study of cognitive development]. In W. Koops & J.J. van der Werf (Eds.), *Overzicht van de ontwikkelingspsychologie: De ontwikkeling van functies en cognitie*. [Outline of developmental psychology: Development of functions and cognition]. Vol 2 Groningen: Wolters-Noordhoff.

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

Wertsch, J.V. (Ed.). (1985). *Culture, communication and cognition. Vygotskian perspectives*. New York: Cambridge University Press.

PSYCHOMETRIC *g* AND THE RACE QUESTION

Arthur R. Jensen

INTRODUCTION

In the United States, the so-called "race question" concerns the conspicuous disparities in educational, economic, and various social criteria between its two largest subpopulations—the majority subpopulation of European descent, hereafter called whites (W), and a minority subpopulation of African descent, called blacks (B). In the 1980 Census, Blacks constituted about 12 percent of the total U.S. population and are projected to reach 15 percent by the year 2020. This chapter focuses on only one of the disparities between these two subpopulations—IQ.

As pointed out in the highly publicized book *The Bell Curve* (Herrnstein & Murray, 1994), whatever it is that IQ tests measure, it is related to nearly all of the educational, economic and social criteria on which whites and blacks, on average, differ, and which, in total, make up what is generally viewed as America's race problem. Herrnstein and Murray's analysis argues strongly that the average B-W disparity in IQ is, to some degree, implicated in all of

Advances in Cognition and Educational Practice, Volume 4, pages 1-23.

ISBN: 0-7623-0105-8

these socially consequential disparities. The main point of contention, however, has not been the existence of the B-W difference in IQ (which is generally acknowledged by social scientists), but the question of whether the mental abilities reflected by IQ are a cause, an effect, or both. This question applies to the whole network of relationships between IQ and the other correlated B-W disparities: scholastic performance, employment, occupational level, income, poverty rate, chronic welfare dependency, infant mortality, health, life expectancy, delinquency and crime.

When IQ is statistically controlled, thereby in effect equating Bs and Ws on IQ, the B-W disparity in many of the social variables listed above is vastly decreased or vanishes completely. This is true for rates of high school and college graduation, income, and occupational level, for example. Therefore most of the average B-W disparity in these social variables, which is often popularly attributed to "racism" in the United States, can be accounted for, statistically if not causally, in terms of the average B-W difference in IQ.

Since *The Bell Curve* juxtaposes race differences in IQ, the heritability of IQ, and the relation of IQ to America's social problems, it has been bashed and trashed by most of the popular media with a vehemence beyond anything ever before accorded to any publication in the social sciences. This is most remarkable, considering that *The Bell Curve* is a scholarly and technically careful work. But such are the consequences of violating the one Great Taboo of modern times.

I will deal here only with the first of these elements—race and IQ. The many interrelated social aspects of both IQ and race are too far-reaching and causally complex for proper consideration in a single chapter. Detailed presentation of the strictly demographic and social aspects (mostly descriptive, scarcely analytical) are provided elsewhere (Jaynes & Williams, 1989; Nichols, 1987; Taylor, 1992). The social variables are most graphically and instructively surveyed in the context of the B-W IQ difference by Herrnstein and Murray (1994).

What IQ Measures

First, it is important to be clear about what IQ tests measure. Although there are many different IQ tests based on different verbal and nonverbal item contents, they all measure mainly *g*, the hypothetical construct originated by Spearman in 1904 and established by means of factor analysis (Spearman, 1927; Jensen, 1987b; Jensen & Weng, 1994). The *g* factor is common to all tests of cognitive ability. Its existence is reflected in the positive correlations among all cognitive tests of every description when such tests are obtained on a representative sample of the general population.

The highest order common factor in a collection of diverse mental tests is *g*. IQ tests are highly loaded on *g*, showing correlations with *g* from .80 to .90. Thus *g* is the single largest component of variance in IQ tests. The

composition of the remaining variance in IQ is divided among group factors (e.g., verbal, spatial, numerical, perceptual), specificity (variance peculiar to a specific test), and error (i.e., unreliability). Since the proportion of IQ variance attributable to each of these non-*g* factors varies from test to test, the average correlation between various IQ tests is only about +.80 (Jensen, 1980a, pp. 313-316), most of which is attributable to the *g* factor.

The physiological basis of individual differences in *g* is not yet precisely known. The fact that *g* has a genetic basis (as shown by its high heritability) and that it is correlated with individual differences in various anatomical and physiological features of the brain argues strongly that *g* reflects the speed and efficiency of information processing via processes in the brain (Vernon, 1993).

On the assumption that *g* is normally distributed in the population, at least within the range of $\pm 2\sigma$ from the mean, the measurement of IQ can be regarded as constituting an interval scale. Since IQ has no true zero point, it is not a ratio scale. An individual's IQ is expressed as a deviation from the mean of a defined (or "normative") population, which is a large, representative sample of the population in the country in which the IQ test is to be used. The normal (Gaussian) distribution of IQ is now universally standardized (or scaled) to $\mu = 100$, $\sigma = 15$ (where μ and σ represent the mean and standard deviation of IQ in the standardization sample, respectively).

Non-psychometric Correlates of Psychometric g

The strongest evidence for the biological reality of psychometric *g* is the fact that *g* is related to other variables and constructs which lie entirely outside the realm of psychometrics and factor analysis and have no connection whatsoever with these methodologies. For example, the degree to which various psychometric tests are *g*-loaded is highly related to their degree of correlation with variables such as the heritability of individual differences in the test scores (Pedersen et al., 1994), the *spouse correlations and various genetic kinship correlations in the test scores, the amount of inbreeding depression* (Jensen, 1983a; Agrawal, et al., 1984) and its converse, *heterosis*, on test performance, *choice reaction time* to visual and auditory stimuli, *inspection time* (i.e., the speed of visual or auditory discrimination), and certain features of the brain's *evoked electrical potentials*. (These studies have been reviewed and referenced in Jensen, 1987b.) No other factor yet extracted from a collection of diverse cognitive tests shows as large or as many correlations with non-psychometric variables as does *g*. It is clear that *g* has as much claim to reality as other theoretical constructs in science.

Psychometric *g* is one of the major constructs in psychology, and one of the oldest and most well-established. Although several different methods of factor analysis have been used by psychometricians to extract the *g* factor from a battery of tests, they all yield highly similar estimates of *g*, as indicated by

congruence coefficients greater than .95 (and averaging .99) between the *g* factors extracted by the different methods (Jensen & Weng, 1994).

The Black-White Difference: Psychometric Evidence

It is important to be clear about two things.

First, we must not conflate two conceptually distinct issues: (1) the observable, empirical fact of the mean black-white difference in mental test score and its many educational and social consequences, and (2) the cause of this observed difference, which has been variously hypothesized as involving, to varying degrees, genetic inheritance and a number of environmental factors, including prenatal and postnatal influences with direct biological effects (e.g., premature birth, nutrition, health care) child-rearing style, educational and cultural deprivation, and societal attitudes toward minority racial groups.

Second, we must be explicit about the specific black and white populations under consideration. The research reviewed in this chapter, like virtually all other psychological (and medical) research on race, is based on samples comprising American-born persons who were self-identified (or parent-identified) as being of a particular racial ancestry. They are so recognized by others as well. This has been called a social definition of race. When individuals so selected are aggregated into groups, this social definition in fact proves to be highly correlated with many biologic or genetic criteria of race. If an observed behavioral difference between certain racial groups, such as the IQ difference, were attributable only to social-cultural factors, as some theories maintain, then the social definition of race should be adequate, and, in fact, should be the only appropriate definition. If two socially defined racial groups are found to differ in some characteristic and one or both groups have some genetic admixture of the other, it means that the racial genetic aspect of the difference, if such exists, has been *under*estimated by comparing the socially defined groups, rather than genetically defined groups.

The population of black Americans, on which all of the findings reported here are based, cannot be regarded the same, racially or genetically, as black (that is, sub-Saharan) Africans. Black Americans are a racially hybrid population, in which about 25 percent of their present gene pool is derived from Caucasian ancestors (Chakraborty et al., 1992; Reed, 1969). The percentage of Caucasian admixture in the black American population varies in different regions, being the lowest (about 10%) in the Deep South, with a positive gradient fanning out to the North and West.

This phenomenon is almost entirely the result of selective emigration of black people from the South, since the black gene pool received Caucasian genes mostly during the period of slavery. Present-day American blacks, therefore, are racially and genetically quite different from their African forebears who were brought to America as slaves. As regards IQ and other tests of general

mental ability, the black population of the United States scores at least as high or higher than black populations in Africa, both in the absolute level of scores and in comparison with the white population in the same locality (Owen, 1992; Zindi, 1994).

The American Black-White IQ Difference

The empirical evidence of the B-W IQ difference is now well-known and uncontentious. A number of key references, taken together, provide a comprehensive account (Eysenck, 1984; Gordon & Bhattacharya, 1994; Herrnstein & Murray, 1994; Jensen, 1973, 1980a, 1994c; Loehlin et al., 1975; Nichols, 1987; Osborne & McGurk, 1982; Shuey, 1966). I will briefly summarize the main points of this evidence, all of which is purely descriptive statistics, no item of which, by itself, should be construed as saying anything about the causes of the observed statistics.

The mean B-W difference in IQ is probably best described in standard deviation (*SD*) units as 1 ± 0.3 *SD* , which is equivalent to 15 ± 5 IQ points when IQ in the total standardization population is scaled to $\mu=100$, $\sigma=15$. The ± 0.3 *SD* allows for the variation among studies attributable to their being based on different IQ tests (which differ in non-*g* composition) and different samples of the black and white populations (which differ by geographical area, age, and other demographic variables). A meta-analysis of virtually all studies (number of studies = 156) of B-W IQ differences published between 1918 and 1990, shows the overall mean of the 156 mean B-W differences of 1.08 *SD*; the 156 mean differences have a $SD = 0.36$. In no study does the black mean IQ exceed the white mean IQ (Herrnstein & Murray, 1994, pp. 276-279).

The *SD* of IQ in the black population is approximately four-fifths of the white *SD* (12 versus 15 IQ points). The degree of overlap of the two distribution is best estimated as mean overlap, or the percentage of the lower-scoring group that falls above the higher group's mean and the percentage of the higher-scoring group that falls below the lower group's mean. Given the black mean IQ = 85, $SD = 12$ and the white mean IQ = 100, $SD = 15$, and assuming a normal distribution of IQ for each population, approximately 11 percent of blacks score above the white mean and approximately 16 percent of whites score below the black mean. The entire range of IQ scores, including the mentally retarded persons and the intellectually gifted, is found within both populations, although with different frequencies.

The mean B-W IQ difference has remained fairly constant, with no clear trend upwards or downwards, since the first large-scale testing of representative samples of blacks and whites in the United States around 1913-1915.

The mean IQ of blacks varies between different parts of the country, being generally lower in the Southeastern states, with an increasing gradient of IQ going toward the Northern and Western states.

Black infants score higher than white infants on developmental scales that depend mainly on sensory-motor abilities. Scores on these infant scales have near-zero correlations with children's IQs at school age, because the IQ predominantly reflects cognitive, rather than sensory-motor, development. Between ages 3 and 4 years, before children normally enter school, the mean B-W IQ difference of about 1 *SD* is fully evident and remains fairly constant from that point on.

Although some IQ tests are intentionally designed to eliminate a sex difference, even tests that were not expressly designed with this aim, such as Raven's Progressive Matrices, show negligible and inconsistent sex differences within the white population (Court, 1983). For as yet unexplained reasons there is a larger, more consistent sex difference in IQ among blacks. Black females average about 4 to 5 IQ points above males. The sex difference in IQ is reflected in scholastic achievement, college admissions and graduation, and occupational status, all of which statistically favor black females (Jensen, 1971).

The B-W difference averages out about the same on verbal and nonverbal IQ tests. The mean B-W difference, however, does vary across many different cognitive tests, each test being composed of homogeneous items such as the subtests of well-known test batteries like the Wechsler Intelligence Scales, the General Aptitude Test Battery (GATB), and the Armed Services Vocational Aptitude Battery (ASVAB).

Variation between tests in the magnitude of the B-W difference is not predictable from the superficial aspects of the tests, such as item content (verbal, nonverbal, or performance), type of test (paper-and-pencil, free response or multiple-choice, individual or group administration, culture loaded or culture reduced). The variable magnitude of the B-W difference on various tests can best be described in terms of factor analysis and the findings of research based on a hypothesis originally suggested by Charles Spearman, which has scientific significance beyond the B-W difference on any particular IQ test.

Spearman's Hypothesis

Referring to an early (1919) study of American B-W differences, in which ten diverse mental tests were used, Spearman (1927) noted that the mean B-W difference "was most marked in just those [tests] which are known to be most saturated with *g*" (p. 379).

I have since referred to Spearman's observation as "Spearman's hypothesis" and have proceeded to test the hypothesis with data from a number of independent studies in which representative samples of blacks and whites were administered six or more diverse mental tests and the correlations among all of the tests were available (Jensen, 1985a, 1985b, 1987a, 1992a, 1993b; Naglieri & Jensen, 1987). It is important that only groups be used that have not been selected on the basis of mental tests or any other *g*-saturated variables (e.g.,

education, occupation) which would seriously restrict the range of *g* relative to other ability factors.

The method for testing Spearman's hypothesis is what I shall term the *method of correlated vectors*. It consists of the following steps: 1. A battery of diverse mental tests is factor analyzed, separately in the black and white samples, by a method that permits the extraction of a general factor, *g*, such as the first factor in a principal factor analysis or the highest order factor in a hierarchical factor analysis.

2. The *g* factor for blacks and whites must be virtually the same factor in both groups, as indicated by a coefficient of congruence greater than .90 (The average congruence coefficient over all of the studies reviewed is close to +.98, which indicates virtual identity of the *g* factor in the black and white groups.) As the *g* factor is essentially the same in both groups, each test's factor loadings in the black and white groups can be averaged, yielding a *column vector* of the average *g* loading of each test. It is incorrect to factor analyze the correlations in the combined groups, because the correlations (and hence the derived factors) would conflate two distinct sources of covariance among the tests—individual differences *within* the groups and the mean difference *between* the groups.

3. The mean B-W difference on each test is expressed in standard score units (i.e., the raw score mean difference divided by the average within-groups *SD*, yielding a column vector of the standardized mean B-W difference on each of the tests).

4. The coefficient of correlation is calculated between the column vector of the tests' *g* loadings and the column vector of the standardized B-W difference. Both the Pearson *r* and Spearman's rank order correlation r_s should be obtained. The assumptions underlying a statistical test of significance of the parametric correlation (Pearson *r*) are not satisfied in the case of these correlated vectors, the elements of which are not statistically independent or normally distributed variates. The nonparametric r_s, however, allows an appropriate test of significance, based on the probability that a purely random rank order of the variates would yield a value of r_s as large or larger than the observed r_s.

5. There is a possible artifact in the correlation (either *r* or r_s) between the vector of *g* loadings and the vector of B-W differences, as the *g* loading and the B-W difference for on each test are both similarly affected by the test's reliability. Hence the correlation between the vectors could be due, wholly or in part, to differences in the reliability of the various tests. This possible artifact must be controlled, either by correcting the *g* loadings and the standardized differences for attenuation (i.e., dividing each by the square root of the test's reliability coefficient) or by partial correlation (i.e., by obtaining the partial correlation between the loadings and the differences with the vector of reliability coefficients partialled out). The resulting correlation (with the effect

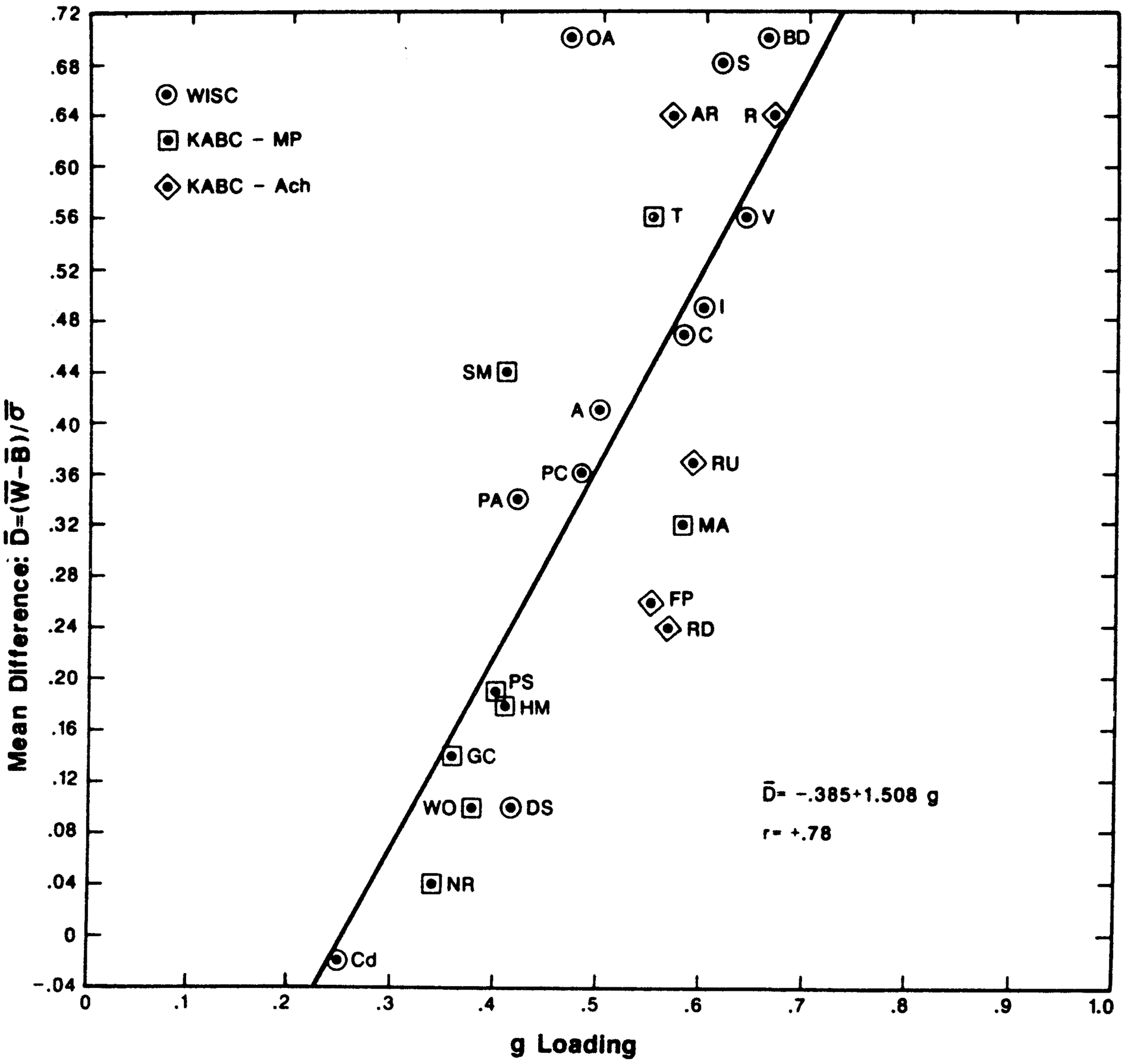
of differences in test reliability controlled) is then used to test Spearman's hypothesis. The r_s is tested for significance by the *t*-test with *n*-2 degrees of freedom, where n is the number of mental tests.

Empirical results of testing Spearman's hypothesis

At present, there are 15 independent data sets, each based on large representative samples of blacks and whites, each of which has taken a battery of anywhere from 6 to 50 diverse cognitive tests. Without exception the results of each study bear out Spearman's hypothesis—a positive correlation between the vector of tests' *g* loadings and the vector of the standardized mean B-W differences on the tests. The average rank-order correlation in the 15 studies is +.60 ($SD = .12$), the median is +.59, and the correlations range from +.36 to +.75. If the significance levels of the 15 independent correlations are combined (Fisher, 1970, p. 99), their joint probability is less than 10^{-10}. There can be no doubt that the magnitude of the B-W difference on a test is directly related to the magnitude of the test's *g* loading. The presence of other factors (independent of *g*) on which there are also B-W differences could only attenuate the test of Spearman's hypothesis, and indeed this is the case.

The lower average socioeconomic status (SES) of blacks and the fact that SES is correlated with IQ within racial groups is commonly suggested as the basis of the B-W difference in IQ. SES differences, however, are more an effect of IQ differences than a cause of them. Matching blacks and whites on SES reduces the mean IQ difference between them only about one-fifth of a *SD* (i.e., from 15 to 12 IQ points). The profile of mean differences between high and low SES groups *within* each racial group on a variety of mental tests, such as the subtests of the WISC-R is markedly different from the profile of B-W differences (Jensen & Reynolds, 1982). The factor structure of SES differences in performance on a variety of psychometric tests is not congruent with the factor structure of the B-W differences.

Evidence that the average SES difference between blacks and whites is not responsible for the phenomenon described by Spearman's hypothesis is shown in a study of black and white school children (grades 4 and 5) matched for age, sex, school attended, and SES (Naglieri & Jensen, 1987). The children were tested on 25 diverse mental tests (the 12 subtests of the Wechsler Intelligence Scale for Children-Revised [WISC-R] and the 13 subtests of the Kaufman Assessment Battery for Children [KABC]), thus providing vectors of 25 *g* loadings and 25 B-W differences. The correlation scatter diagram (with Pearson $r = +.78$, rank-order correlation = +.75, p .01) between the vectors is shown in Figure 1. The correlation is not higher in this study (and in other studies), in part, because the data reveal that Spearman's hypothesis cannot be accepted in its pure form, but needs to be modified. Although the B-W difference is largely a difference in *g*, as stated by Speartman's hypothesis,



Key: WISC-R subtests: I-Comprehension, DS-Digit Span, PC-Picture Completion, PA-Picture Arrangement, BD-Block Design, OA-Object Assembly, Cd Coding. K-ABC subtests: HM-Hand Movements, GC-Gestalt Closure, NR-Number Recall, T-Triangles, WO-World Order, MA-Matrix Analogies, SM-Spatial Memory, PS-Photo Series, FP-Faces and Places, Ar-Arithmetic, R-Riddles, RD-Reading/Decoding, RU-Reading/Understanding

Figure 1. Mean B-W differences (expressed in units of the average within-groups standard deviation) on WISC-R and K-ABC subtest as a functions of the subtests' loadings on g

for certain tests part of the difference is also attributable to either one of two group factors, independently of *g*. Two classes of tests, one loaded on a *spatial* factor and the other loaded on a *short-term rote memory* factor, fall rather consistently above and below the regression line, respectively. The B-W difference is larger on spatial tests than would be predicted by just their *g* loadings, and smaller on tests involving short-term memory than would be predicted by their *g* loadings. In comparing the B-W difference on tests that

are about equally loaded on either a spatial or a memory factor, the relative magnitude of the B-W difference is predictable from the tests' *g* loadings alone. For example, the forward digit span (FDS) and backward digit span (BDS) subtests of the WISC have no loadings on a spatial factor and similar loadings on a short-term memory factor, but the *g* loading of BDS is approximately twice the *g* loading of FDS, and the standardized mean B-W difference is approximately twice as great on BDS as on FDS (Jensen & Figueroa, 1975).

The regression equation in Figure 1 predicts that on a hypothetical pure test of *g* (i.e., a *g* loading of 1.0) the mean standardized B-W difference (for these particular samples) would be 1.12. This is an underestimate of the population value, because the groups in this study, having been matched on SES, differ by only .73 on the WISC-R Full Scale IQ, as compared to the 1.14 difference in the representative national standardization sample of the WISC-R (Jensen & Reynolds, 1982). The same kind of correlated vectors analysis performed on the WISC-R standardization sample (with 305 blacks and 1,870 whites) predicts a B-W difference of 1.3 for a hypothetical pure test of *g*.

The Theoretical Importance of Spearman's Hypothesis

The empirical fact that the B-W difference on each cognitive test is highly related to that tests' *g* loading is a phenomenon that calls for explanation. Why is the B-W difference on tests so greatly reduced when the *g* component is removed (Reynolds & Jensen, 1983)? The reason cannot be that *g* is the overwhelming source of variance in cognitive tests, since in a diverse battery it usually accounts for less of the total variance than is accounted for by various group factors (e.g., verbal, spatial, numerical) and test-specific variance. Yet it is *g*, rather than the non-*g* components of test score variance, that best predicts the magnitude of the B-W difference. The B-W difference is essentially a difference in *g*. Spatial tests increase the difference slightly, over and above the difference in *g*, and short-term memory tests (e.g., digit span, coding, rote-learning) slightly subtract from the difference due to *g*.

It is important to distinguish between the terms *culture loaded* and *culture biased* as applied to tests (Jensen, 1980a, Chapter 9). To say that a test is culture loaded simply means that the item content of the test contains specific knowledge and skills acquired in a particular culture. A culture loaded test may or may not be culture-biased for a particular population. This is a question that must be empirically tested. A test is biased with respect to the difference between two groups (regardless of the size or direction of the mean difference between the groups) if it behaves differently in each group in its essential psychometric features, such as internal consistency reliability, factor composition, construct validity, and practical predictive validity for educational and work-related variables.

What is commonly called culture bias in tests, to the extent that it exists at all, seems to reside in the non-*g* components of the test, for the B-W difference on any test is better predicted by the test's *g* loading (regardless of its cultural content) than by its cultural content (regardless of its *g* loading). Reducing the culture-specific content of IQ test items does not reduce the B-W difference and in fact *increases* the B-W difference (Jensen & McGurk, 1987). Every possible kind of test bias that has been hypothesized to explain the mean B-W disparity in mental test scores that has been empirically investigated has been ruled out as a cause of the observed B-W difference (Jensen, 1980a, 1984b). Standard psychometric tests measure the same factors in blacks and whites; and they have the same validity for predicting a wide variety of educational, occupational, and job-performance criteria. The substantiation of Spearman's hypothesis is evidence that our present tests also have the same *g* construct validity for blacks and whites.

The cause of the B-W difference must therefore be sought in studies of the nature of *g* itself. If the *g* factor were found not to be the same factor for blacks and whites in terms of its loadings on our present psychometric tests, then examining the nature of *g* would not be a potentially fruitful approach to understanding the nature of the B-W differences on psychometric tests. It so happens, however, that the *g* that predicts the mean B-W difference across various tests is the very same *g* that emerges from the factor analysis of families' (full siblings) mean test scores *within* either the white group or the black group, and the very same *g* that emerges from the factor analysis of the test scores obtained by individuals *within* the same family (i.e., full siblings reared together by their biological parents). In brief, one and the same *g* exists (1) *between* racial groups, (2) *between* families *within* each racial group, and (3) *between* individuals *within* the same family (Jensen, 1980b).

SES differences and cultural differences are generally hypothesized to be more potent causes of differences between families than *within* families. (The rationale and methodology for between and *within* family analyses are explained in Jensen, 1980b.) If SES and cultural differences per se were important determinants of test scores, it is hard to imagine why a battery of highly diverse tests would reflect the same *g* when the basic elements in the factor analysis are family means and when the basic elements in the factor analysis are *differences* between siblings *within* families.

Using the same methodology by which a column vector of various tests' *g* loadings is correlated with the column vector of the mean B-W differences on the tests, the vector of tests' *g* loadings can be correlated with the vector of the tests' correlations with some other variable that may hold a clue to the nature of the B-W difference. The nonpsychometric correlates of *g* may suggest potentially fruitful hypotheses as to the nature of the B-W difference. Applying the method of correlated vectors suggests that something biological is implicated. The vector of various tests' heritability coefficients, of tests' index

of inbreeding depression, of tests' correlations with average evoked cortical potentials, tests' correlations with head size (as a proxy for brain size), and tests' correlations with various measures of reaction time in elementary cognitive tasks are each highly correlated with the vector of the tests' *g* loadings. It is virtually impossible to explain all these correlated vectors without positing a physical basis for *g*. It simply isn't probable that social-cultural factors alone could produce all of these correlated vectors. Further, it is highly improbable that these vectors of biological indices would be correlated with the vector of B-W differences simply by chance. Some causal connection is far more likely.

Indeed, some of these vectors have already been directly correlated with the vector of B-W differences on the same set of tests.

Inbreeding Depression. The vector of *g* loadings of the WISC subtests predicts the vector of the degree of inbreeding depression manifested on reach subtest, and the vector of the magnitude of inbreeding depression on each of the subtests is significantly correlated with the vector of B-W difference on each of the subtests (Rushton, 1989).

Reaction Time. Research on reaction time (RT) to elementary cognitive tasks has shown that RT variables (such as the individual's median RT and the *SD* of RTs over a number of trials), which measure the speed and efficiency of information processing, are negatively correlated with *g* (Jensen, 1982, 1987b, 1988, 1992b, 1993a, 1994a, Vernon, 1987). (The correlation is negative, because shorter, that is, faster, RT is associated with higher scores on *g*-loaded tests.) Also, various tests' *g* loadings predict the magnitude of each test's correlation with RT. When *g* is partialled out of the correlation, this correlation is reduced almost to zero. In a study (Jensen, 1993b) of B-W differences in 24 RT variables derived from RT tasks of varying complexity and that are so simple that the subjects (children in grades 4, 5, and 6) could perform the most difficult of them in less than 2 seconds, the vector of *g* loadings of the 24 RT variables was correlated with the vector of B-W differences on each of the 24 RT variables (Pearson $r = +.82$, rank-order correlation $= +.70$, $p < .01$). That is, blacks generally had longer RTs and greater inter-trial variability in RT on tasks for which RT was more *g* loaded. Thus Spearman's hypothesis is confirmed for RT tasks as well as for conventional psychometric tests.

A B-W difference in motivation is not a plausible explanation of these results. RT was measured as the time between the onset of the reaction stimulus (a green light going "on") and the subject's removing the index finger from a "home" button in response to the reaction stimulus. The interval between lifting the finger from the home button and pressing another button (at a distance of 6 inches from the home button) that turns off the reaction stimulus is a second measure called movement time (MT). Blacks had faster MT than whites, but MT was not significantly correlated with *g* within either group.

It is implausible that motivation or effort would work in opposite directions for blacks and whites during the brief RT-MT interval (typically less than 1 second). In fact, subjects generally perceive both the RT and MT aspects of the task as simply making a single ballistic response to the onset of the reaction stimulus, taking their finger off the "home" button and pressing another button to turn off the reaction stimulus as quickly as they can.

Head Size and Brain Size. These variables are correlated with each other (about +.50), and each is correlated with IQ (+.15 to +.20 for external head size, +.35 to +.40 for brain size) (Andreasen et al., 1993; Jensen & Sinha, 1993; Rushton & Ankney, in press, is an exhaustive review). External head-size is merely an attenuated proxy for brain size. Statistically controlling body size (height, weight, total surface area of the body) scarcely reduces the head/brain-size correlation with IQ. Moreover, the head-size × IQ correlation exists *within* families as well as between families, in both the white and the black populations (Jensen & Johnson, 1994). In other words, in a given family, the sibling with the larger head, on average, has a higher IQ than the sibling with a smaller head (with head-size adjusted for age). Families (i.e., full siblings) with a larger average head-size tend to be higher in average IQ than families with a smaller average head-size. The *within* family correlation between head-size and IQ indicates an intrinsic or functional relation between these variables and is consistent with *pleiotropy*, that is, one and the same gene (or set of genes) determines more than one distinct phenotypic character.

There are also racial differences in both head-size and brain size (measured either as weight or volume), which are not appreciably diminished when overall body size is controlled (Jensen & Sinha, 1993; Rushton & Ankney, in press). The average B-W difference in autopsy brain weight (for American whites and blacks) is about 100 grams, equivalent to about 0.8 *SD* (based on the average within-group *SD*).

Head-size and g. Head size (measured as length, width, and circumference), as a proxy for brain size, was shown to be directly related to psychometric *g* by the method of correlated vectors (Jensen, 1994b). The *g* factor was extracted by a hierarchical (Schmid-Leiman) factor analysis from a battery of 17 diverse mental tests separately in large samples of black and of white students (ages 12 to 20). Since the *g* factor was virtually identical in the black and white samples (congruence coefficient of +.99), the *g* loadings on each of the 17 tests were averaged across the black and white groups to yield a single column vector of the 17 tests' *g* loadings.

The multiple correlation (R) of the head-size measurements (length, width, circumference) with each of the tests was obtained. (The mean R was 0.21, ranging from 0.10 to 0.30 for the various tests. All of the correlations are significant at $p < .05$.) The single column vector of the 17 tests' *g* loadings

was correlated with the column vector of the tests' correlations with head-size. The rank-order correlation was +.62, $p < .01$. The vector of the tests' loadings on the spatial factor (independent of *g*) correlated + .27 with the head-size vector; the verbal factor (sans *g*) correlated .00. Head-size is reflected mostly by tests that are highly loaded on both *g* and spatial ability. The multiple correlation of the vectors of *g* factor loadings and spatial factor loadings with the vector of head-size correlations was .71, $p < .01$.

Relation of Head-size to Spearman's Hypothesis

The column vector of the combined *g* loadings and spatial loadings on each of the 17 tests was rank correlated +.81 with the corresponding column vector of standardized B-W differences on each of the tests (controlled for differences in test reliability). This bears out the revised Spearman hypothesis, namely, that the magnitude of the B-W difference on various psychometric tests is essentially a function of the test's loading on *g*, but is slightly increased on tests that are loaded on the spatial factor.

If head (or brain) size is one of the causal variables mediating the correlation between the *g* + spatial vector and the B-W differences on the various tests, we would predict a correlation between the vector of head-size correlations with the *g* + spatial factor loadings and the vector of B-W differences. The rank correlation (controlled for differences in test reliability) was +.72, $p < .01$. That is, the magnitude of the B-W difference on various tests was predicted (with a validity coefficient of .72) by the tests' correlations with head-size.

Tests loaded on *g* and spatial factors share a common source of variance with both head (or brain) size and the mean B-W difference in test scores. This is consistent with the finding (in an independent study based on very large samples) that black and white children who are matched on IQ do not differ in head size, although the total white and black samples from which the matched groups were drawn differed about 1σ in IQ and differed significantly in head size (Jensen & Johnson, 1994). This result implies that only the component of B-W difference in brain-size that is associated with IQ differs between the black and white groups. When black and white groups are equated for brain-size, they still show a mean difference (though reduced) in IQ, suggesting that brain-size is only one among many variables associated with IQ.

Myopia and IQ. Many studies (reviewed in Cohn et al., 1988, and in Jensen & Sinha, 1993) have shown that myopia is positively related to IQ. Furthermore, the correlation between IQ and myopia (measured as a continuous variable) exists *within* families; that is, the sibling with the higher IQ, on average, is more myopic than his or her less intellectually-gifted sibling (Cohn, et al., 1988). Such a within-family correlation between highly heritable

variables as IQ and myopia is consistent with pleiotropy, which implies an intrinsic genetic relationship between the correlated variables.

If IQ (which is highly *g*-loaded) and myopia are pleiotropic, one would expect a lower incidence of myopia in persons of African ancestry than in persons of European ancestry. This in fact is what is found. When the entire population of young men in the United States was examined by the Selective Service system, in the 1940s, during World War II, the incidence of myopia in whites and blacks was 19.3 percent and 3.4 percent, respectively (Post, 1982). The white rejection rate for myopia was over three times that for blacks. (Assuming a normal distribution of refraction error on the myopia-emmetropia dimension within each racial group, the average difference between the means of the black and white difference on this dimension is about one standard deviation.) It is noteworthy that the incidence of myopia is highest among Asians and Jews—groups that score above the mean IQ in the general population (Herrnstein & Murray, 1994, pp. 272-301).

Causal Factors of the Black-White Difference in g

The findings reviewed above are difficult to explain in purely socio-cultural terms. No non-biological environmental factor has yet been suggested that can account for these observed relationships of *g* to biological variables and to race. These findings suggest that biological factors are involved in the B-W difference in *g*. The biological factors could have both environmental and genetic aspects, but their relative contributions have not been determined. Probably some part of the IQ difference is attributable to those biological environmental factors on which the black and white populations, on average, are known to differ, such as prenatal health care, maternal age and health, prematurity and low birth weight, prevalence of breast-feeding and child nutrition, and inoculation against childhood diseases. Genetic factors should not be rejected out of hand or dogmatically condemned as "scientific racism, "pseudoscience," or the like.

At present, however, there is hardly a consensus of opinion about the probable role of genetic factors in the B-W IQ difference. A questionnaire survey (Snyderman & Rothman, 1988) of 661 experts, mostly in differential psychology, psychometrics, and behavioral genetics, found the following response percentages to the question:

Which of the following best characterizes your *opinion* of the heritability of the black-white difference in IQ?

15	The difference is entirely due to environmental variation.
1	The difference is entirely due to genetic variation.
45	The difference is a product of both genetic and environmental variation.

24 The data are insufficient to support any reasonable opinion.
14 NQ [Does not feel qualified to answer question.]

Of course, mere opinions, even of experts, are not to be mistaken for empirical evidence. But one may ask what evidence would be consistent with the predominant opinion ("a product of both genetic and environmental variation"), expressed by 53 percent of those who did not disqualify themselves. A thorough consideration of all of the arguments that have been put forth, of course, would go far beyond the allotted length of this chapter. Therefore, I can mention only briefly the most salient items of evidence that make it reasonable to investigate further the role of genetic, as well as environmental factors, in the causation of the well-established B-W difference in *g*.

Evolution. The broadest consideration is the theory of evolution by natural selection, which explains the origin of genetic differences between subspecies, or races, that have been geographically separated for hundreds of generations in markedly differing environments (Rushton, 1995). These influences are reflected in the many physical differences among various races. Most polymorphic traits that vary within racial groups also vary between racial groups. It seems extremely implausible that such a highly complex system as the brain—its anatomy, physiology, and polygenic behavioral correlates, including cognitive abilities—would be the only system exempt from variation between different racial populations. If we grant that perfect genetic equality for a given trait among all known racial populations is exceedingly improbable, it is also unlikely that, under natural conditions, genotypes and phenotypes for any characteristic are *negatively* correlated. A negative phenotype-genotype correlation seems especially unlikely for a trait, such as *g*, with high heritability and a wide range of phenotypic expression. (If the broad heritability of phenotypic adult IQ *within* a racial group is .70, the within-group coefficient of correlation between phenotypes and genotypes would be the square root of .70, that is, +.84.) The largest study of both white and black twins found approximately equal heritability for IQ *within* each racial group (Osborne, 1980).

The idea that the observed mean differences *between* populations are attributable to the same genetic and nongenetic factors (though perhaps differing in relative effects) that cause individual differences in polygenic traits *within* populations can be called the "default" position. It is the simplest hypothesis, and, without strong evidence to the contrary, the most plausible. A purely nongenetic model demands the demonstration of either extreme between-groups environmental differences or some yet unspecified environmental influence unique to one or the other group, which I have referred to as Factor X (Jensen, 1873, pp. 164-169). It is the default hypothesis, not the null hypothesis (which posits a zero genetic difference in *g* between populations), that should be the target of empirical tests.

Between-groups Heritability. Theoretically the genetic and environmental causes of the mean phenotypic difference between populations can be different from the genetic and environmental causes of individual differences *within* each population. However, the claim that the causal factors do in fact differ *between* populations requires the identification of some genetic or nongenetic cause(s) that simultaneously satisfies three criteria—(1) an independently demonstrable effect of the causal factor on the phenotype, (2) a sufficient magnitude of the effect to account for that part of the between-populations variance attributed to the causal factor, and (3) a difference between populations in the incidence of the causal factor. If it is hypothesized that all the causal factors in the between-populations difference are purely environmental (i.e., nongenetic), the above requirements hold for any and all of the specific environmental factors claimed to be causal.

The kinds of social-psychological variables most commonly hypothesized as the main causes of the W-B IQ difference, such as differences in educational, cultural, socioeconomic background, and child-rearing practices, are environmental variables, which, to the extent that they affect IQ at all, constitute part of the between-families (BF) environmental variance in IQ.

The remaining environmental variance is the *within*-family (WF) component, which consists largely of random micro-environmental effects, acting mostly at the biological level (Jensen, 1997). One of the new important discoveries of behavioral genetics is that the component of BF environmental variance in IQ, though evident in childhood, gradually diminishes after puberty, reducing to virtually zero by late adolescence (Plomin & Daniels, 1987). Therefore, social-cultural BF differences seem an unpromising explanation for the B-W IQ difference, especially in the common American culture. The majority of black and white children share a common culture, attend the same schools, watch the same TV programs, play with the same toys, and consume nutritionally similar diets.

The most popular causal hypotheses invoke SES differences, language deprivation, and educational inequality. They have been examined in a great many studies and simply fail to explain the B-W difference (Jensen, 1973).

A host of potential microenvironmental biological effects during gestation and infancy that are adverse to later mental development are plausibly more prevalent in a large segment of the black population. They have not yet been studied adequately to be proven to be causal factors. I suspect that they constitute most of whatever part of the B-W difference in *g* is not attributable to genetic factors per se. Two findings in particular lead me to believe this: (1) the gradual but clear-cut *within-family* decline of IQ with age in black children reared in poor environments (Jensen, 1977) and (2) the sex difference in IQ and scholastic achievement (female $>$ male) in the black population (Jensen, 1971).

Although the preponderance of evidence supports the hypothesis that both genetic and nongenetic factors contribute to the between-groups difference, it does not provide a precise estimate of the between-groups heritability. However, certain types of research may at least point to a reasonable approximation.

The default hypothesis of the phenotypic B-W difference in *g*-loaded IQ would predict that if a representative sample of the black population were reared by white adoptive parents in the typical family environment of advantaged whites, the blacks' level of *g*, as measured by IQ, should still differ from the white average IQ by an amount proportional to the genetic difference between the groups, assuming no adverse prenatal effects for the black adoptees. On the other hand, the hypothesis that the IQ deficit of the average black merely reflects the lack of acculturation to the white middle-class environment would predict that the black adoptees, after growing up in a white home, should have an average IQ on a par with whites.

These opposing hypotheses were tested in a natural quasi-experiment known as the Minnesota Transracial Adoption Study (Scarr & Weinberg, 1976; Weinberg, Scarr, & Waldman, 1992). Black children ($N = 29$) with two blacks parents (B/B) and inter-racial children ($N = 109$) with a black father and a white mother (B/W), who had been given up for foster care or adoption shortly after birth, were adopted by white parents at an average age of about one year. The adoptive parents were mostly college-educated and upper-middle class. Also included in the study were the biological children of the adoptive parents ($N = 118$) and white adoptees, W/W ($N = 18$). All the children were tested on the Wechsler scales at age 7 (in 1975) and at age 18 (in 1986). The mean IQ of each group is shown in Table 1.

It is evident that these data are more consistent with the default hypothesis than with a strictly environmental hypothesis:

Table 1. IQ Mean and Standard Deviation of Groups in the Minnesota Transracial Adoption Study

	1975			*1986*	
Group	*N*[a]	*Mean*	*SD*	*Mean*	*SD*
Adoptive Father	74	*121.7*	*9.5*	*117.1*	*11.5*
Adoptive Mother	84	*119.1*	*9.7*	*113.6*	*10.5*
Biological Offspring	104	*116.4*	*13.5*	*109.4*	*13.5*
Adopted Interracial (B/W)	55	*109.5*	*11.9*	*98.5*	*10.6*
Adopted Black (B/B)	21	*95.4*	*13.3*	*89.4*	*11.7*
Adopted White (W/W)	16	*117.6*	*11.3*	*105.6*	*14.9*

Note: [a] The number of individuals tested both in 1975 and 1986.

1. Both at age 7 and at age 18 the white adoptees' differ much less from the adoptive parents' biological offspring than do either the interracial (B/W) or black (B/B) adoptees.
2. The mean IQ of the B/W adoptees is midway between the means for the B/B and W/W adoptees.
3. A 16-point IQ difference separates the means of the W/W and B/B at age 18—a difference nearly the same as the mean B-W difference (15.8 IQ points) in the Wechsler national standardization data and the average of 156 published studies of the B-W difference (16.2 IQ points). (There are no normative data with which to compare the interracial [B/W] adoptees.)
4. The B/B adoptees at age 18 have approximately the same mean IQ as that of blacks (reared from birth in black families) in those regions in which the adoptees were born (Northeastern and Northcentral United States).

Although the authors of this hardly favor a genetic interpretation of their data, others have argued in explicit detail how and why the data implicate genetic factors in the B-W IQ difference (Levin, 1994; Lynn, 1994). On the basis of these data, in fact, Levin (1994) tried to estimate the *between*-groups heritability for IQ as .59, .66, and .70, depending on different reasonable assumptions about the quality of the adoptive environment. Although these between-group heritability values are by no means definitive, given the small sample sizes and Levin's unconventional method of heritability estimation, they are probably the best estimates that could be gleaned from this set of data. The estimated values of the between-groups heritability, such as they are, fall well within the mid-range of heritability coefficients estimated from twin studies for individual differences in IQ (Bouchard et al., 1990). The position argued by Levin and by Lynn is not at all contradicted by the Minnesota Study authors' reply (Waldman, Weinberg, & Scarr, 1994). Their study, even if it proves nothing else, clearly refutes the cultural difference explanation of the B-W IQ difference.

A New Methodology for Studying Group Differences

In animal research the genetic and environmental components of the difference between two strains can be determined by the direct methods of experimental genetics, such as random assignment to different environments and cross-breeding of strains. These methods obviously are neither feasible nor ethically permissible with humans. Modern techniques in molecular genetics now make it possible to identify specific trait-related genes when a large phenotypic effect is associated with single gene locus. In the case of highly polygenic traits, for which the variance is attributable to a large number of

genes each having a relatively small effect on the trait, the analytic task is almost forbidding—like looking for a needle in a haystack. Yet this is now being attempted (Plomin & Neiderhiser, 1991), but it is apt to be a good many years, if ever, before the results of such research can be informative as to any specific genes being involved in racial population differences.

These technical problems, however, do not rule out the possibility of testing what I have termed the default hypothesis of racial group differences, that is, that the sources of variation between groups are the same as for individual differences within each group.

Sophisticated methods of biometrical genetics have been developed that are capable of rigorously testing the default hypothesis and estimating the heritability of the difference between group means. This class of methods is now known as the biometric decomposition of phenotypic means (Dolan, 1992; Dolan et al., 1994; Turkheimer, 1991). It can provide both a statistical test of the default hypothesis (namely, that genetic and nongenetic factors contribute to the mean difference between groups) and a "best estimate" (in terms of Maximum Likelihood) of the between-groups heritability, that is, the proportions of genetic and environmental variance between the group means. Because this method is mathematically and statistically complex, a technical exposition is beyond the scope of this chapter.

The biometrical decomposition of phenotypic means requires special data sets that meet certain preliminary requirements: (1) A number of different tests, not just one test, must be used to measure the construct of interest (e.g., *g*). (2) The battery of tests must measure the same construct in each group, as indicated by a high coefficient of congruence between the general factor extracted from the test battery within each group. (3) The test data obtained within each group must be suitable for heritability analysis (i.e., the estimation of genetic and environmental variance components) *within* each group, for example, requiring subject samples comprising many pairs of monozygotic and dizygotic twins or other types of kinship data that would allow biometrical analysis. (4) The construct of interest should have approximately the same heritability coefficient within each group.

Given these four conditions, and statistically controlling any extraneous variables that could confound the results (such as differences in age, sex, and SES), this biometrical method is capable of statistically rejecting (or not rejecting) any one of the prevailing hypotheses regarding the mean group difference—that it is wholly genetic (G), wholly environmental (E), or both G and E. If the G and the E models can each be rejected at a high level of significance, and the G + E model cannot be rejected, it would support the default hypothesis. The method could also provide best-fit estimates of the proportions of the between-groups variance attributable to G and E. Assuming adequate sample size and replicated studies, this methodology, properly applied, should provide the best answer to the question of group differences

in *g* that scientists are realistically capable of obtaining in the immediate future. If the default hypothesis were borne out by rejecting the alternative hypotheses at a high level of significance, the argument that some unidentified nongenetic Factor X, rather than polygenes, explained the mean phenotypic difference between the groups could be ruled out as a violation of Occam's razor.

REFERENCES

Agrawal, N., Sinha, S.N., & Jensen, A.R. (1984). Effects of inbreeding on Raven Matrices. *Behavior Genetics, 14*, 579-585.

Andreasen, N.C., Flaum, M., Swayze II, V., O'Leary, D.S., Alliger, R., Cohen, G., Ehrhardt, J., & Youh, W.T.C. (1993). Intelligence and brain structure in normal individuals. *American Journal of Psychiatry, 150*, 130-134.

Chakraborty, R., Kamboh, M. I., Nwankwo, M., & Ferrell, R. E. (1992). Caucasian genes in American blacks: New data. *American Journal of Human Genetics, 50*, 145-155.

Cohn, S.J., Cohn, C.M.G., & Jensen, A.R. (1988). Myopia and intelligence: A pleiotropic relationship? *Human Genetics, 80*, 53-58.

Court, J.H. (1983). Sex differences in performance on Raven's Progressive Matrices: A review. *The Alberta Journal of Educational Research, 29*, 54-74.

Dolan, C.V. (1992). *Biometric decomposition of phenotypic means in human samples*. Amsterdam: University of Amsterdam.

Dolan, C.V., Molenaar, P.C.M., & Boomsma (1994). Simultaneous genetic analysis of means and covariance structure: Pearson-Lawley selectionrules. *Behavior Genetics, 24*, 17-24.

Eysenck, H. J. (1984). The effect of race on human abilities and mental test scores. In C. R. Reynolds & R. T. Brown (Eds.), *Perpectives on bias in mental testing* (pp. 249-291). New York: Plenum.

Fisher, R.A. (1970). *Statistical methods for research workers*. 14th ed. New York: Hafner Press.

Gordon, E.W., & Bhattacharya, M. (1994). Race and intelligence. In R.J. Sternberg (Ed.), *Encyclopedia of human intelligence* (Vol. 2, pp. 889-899). New York: Macmillan.

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

Jaynes, G. D., & Williams, R. M., Jr. (Eds.) (1989). *A common destiny: Blacks and American society*. Washington, DC: National Academy Press.

Jensen, A. R. (1971). The race sex ability interaction. In R. Cancro (Ed.), *Contributions to intelligence* (pp. 107-161). New York: Grune & Stratton.

Jensen, A. R. (1973). *Educability and group differences*. London: Methuen.

Jensen, A.R. ((1977). Cumulative deficit in IQ of blacks in the deep south. *Developmental Psychology, 13*, 184-191.

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* (pp. 93-132). New York: Springer.

Jensen, A.R. (1983a). Effects of inbreeding on mental-ability factors. *Personality and Individual Differences, 4*, 71-87.

Jensen, A.R. (1984b). Test bias: Concepts and criticisms. In C.R. Reynolds & R.T. Brown (Eds.), *Perspectives on bias in mental testing* (pp. 507-586). New York: Plenum.

Jensen, A.R. (1985a). The nature of the black-white difference on various psychometric tests: Spearman's hypothesis. *Behavioral and Brain Sciences*, 8, 193-219.

Jensen, A.R. (1985b). The black-white difference in g: A phenomenon in search of a theory. *Behavioral and Brain Sciences, 8*, 246-263.

Jensen, A.R. (1987a). Further evidence for Spearman's hypothesis concerning the black-white differences on psychometric tests. *Behavioral and Brain Sciences, 10*, 512-519.

Jensen, A.R. (1987b). The *g* beyond factor analysis. In R.R. Ronning, J.A. Glover & J.C. Conoley (Eds.), *The influence of cognitive psychology on testing* (pp. 87-142). Hillsdale, NJ: Erlbaum.

Jensen, A.R. (1988). Speed of information processing and population differences. In S.H. Irvine & J.W. Berry (Eds.), *Human abilities in cultural context* (pp. 105-145). Cambridge: Cambridge University Press.

Jensen, A.R. (1992a). Spearman's hypothesis: Methodology and evidence. *Multivariate Behavioral Research, 27*, 225-233.

Jensen, A.R. (1992b). The importance of intraindividual variability in reaction time. *Personality and Individual Differences, 13*, 869-882.

Jensen, A.R. (1993a). Spearman's g: Links between psychometrics and biology. *Annals of the New York Academy of Sciences, 702*, 103-131.

Jensen, A.R. (1993b). Spearman's hypothesis tested with chronometric information-processing tasks. *Intelligence, 17*, 47-77.

Jensen, A.R. (1994a). Reaction time. In R.J. Corsini (Ed.), *Encyclopedia of psychology* (Vol. 3, pp. 282-285). New York: Wiley.

Jensen, A.R. (1994b). Psychometric *g* related to differences in head size. *Personality and Individual Differences, 17*, 597-606.

Jensen, A.R. (1994c). Race and IQ scores. In R.J. Sternberg (Ed.), *Encyclopedia of human intelligence* (Vol. 2, pp. 899-907). New York: Macmillan.

Jensen, A.R. (1997). The puzzle of nongenetic variance. In R.J. Sternberg & E.L. Grigorenko (Eds.) *Intelligence, heredity and environment* (pp. 107-124). Cambridge: Cambridge University Press.

Jensen, A.R., & Figueroa, R.A. (1975). Forward and backward digit span interaction with race and IQ: Predictions from Jensen's theory. *Journal of Educational Psychology, 67*, 882-893.

Jensen, A.R., & Johnson, F.W. (1994). Race and sex differences in head size and IQ. *Intelligence, 18*, 309-333.

Jensen, A.R., & McGurk, F.C.J. (1987). Black-white bias in 'cultural' and 'noncultural' test items. *Personality and Individual Differences, 8*, 295-301.

Jensen, A.R., & Reynolds, C.R. (1982). Race, social class and ability patterns on the WISC-R. Personality and Individual Differences, 3, 423-438.

Jensen, A.R., & Sinha, S.N. (1993). Physical correlates of human intelligence. In P.A. Vernon (Ed.), *Biological approaches to the study of human intellligence* (pp. 139-242). Norwood, NJ: Ablex.

Jensen, A.R., & Weng, L-J. (1994). What is a good g? *Intelligence, 18*, 231-258.

Levin, M. (1994). Comment on the Minnesota Transracial Adoption Study. *Intelligence, 19*, 13-20.

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

Lynn, R. (1994). Some reinterpretations of the Minnesota Transracial Adoption Study. *Intelligence, 19*, 21-28.

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

Nichols, R. C. (1987). Racial differences in intelligence. In S. Modgil & C. Modgil (Eds.), *Arthur Jensen: Consensus and controversy* (pp. 213-220). New York: The Falmer Press.

Osborne, R. T. (1980). *Twins, Black and White*. Athens, Georgia: Foundation for Human Understanding.

Osborne, R. T., & McGurk, F. C. J. (1982). *The testing of Negro intelligence, Vol. 2.* Athens, Georgia: Foundation for Human Understanding.

Owen, K. (1992). The suitability of Raven Standard Progressive Matrices for various groups in South Africa. *Personality and Individual Differences,13,* 149-159.

Pedersen, N.L., Plomin, R., & McClearn, G.E. (1994). Is there G beyond g? (Is there genetic influence on specific cognitive abilities independent of genetic influence on general cognitive ability?). *Intelligence, 18*, 133-143.

Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from one another? *Behavioral and Brain Sciences, 10*, 1-16.

Plomin, R., & Neiderhiser, J.M. (1992). Quantitative genetics, molecular genetics, and intelligence. *Intelligence, 15*, 369-387.

Post, R.H. (1962). Population differences in visual acuity: A review, with speculative notes on selection relaxation. *Social Biology, 9*, 189-212.

Reed, T. E. (1969). Caucasian genes in American Negroes. *Science, 165,* 762-768.

Smith, G.A., & Stanley, G. (1987). Comparing subtest profiles of *g* loadings and correlations with RT measures. *Intelligence, 11,* 291-298.

Reynolds, C.R., & Jensen, A.R. (1983). WISC-R subscale patterns of abilities of blacks and whites matched on Full Scale IQ. *Journal of Educational Psychology, 75*, 207-214.

Rushton, J.P. (1989). Japanese inbreeding depression scores: Predictors of cognitive differences between blacks and whites. *Intelligence, 13*, 43-51.

Rushton, J.P. (1995). *Race, evolution, and behavior.* New Brunswick, NJ: Transaction

Rushton, J.P., & Ankney. (1996). Brain size and cognitive ability: Correlational relationships. *Psychonomic Bulletin and Review, 3*, 21-36.

Scarr, S., & Weinberg, R.A. (1976). IQ test performance of black children adopted by white families. *American Psychologist, 31*, 726-739.

Shuey, A. M. (1966). *The testing of Negro intelligence*, 2nd ed. New York: Social Science Press.

Snyderman, M., & Rothman, S. (1988). *The IQ controversy: The media and public policy*. New Brunswick, NJ: Transaction Books.

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

Taylor, J. (1992). *Paved with good intentions: The failure of race relations in contemporary America*. New York: Carroll & Graf.

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

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

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

Waldman, I.D., Weinberg, R.A., & Scarr, S. (1994). Racial-group differences in IQ in the Minnesota Transracial Adoption Study: A reply to Levin and Lynn. *Intelligence, 19*, 29-44.

Weinberg, R.A., Scarr, S., & Waldman, I.D. (1992). The Minnesota Transracial Adoption Study: A follow-up of IQ test performance in adolescence. *Intelligence, 16*, 117-135.

Zindi, F. (1994). Differences in psychometric performance. *The Psychologist, 7*, 549-552.

RACE AND SEX DIFFERENCES IN BRAIN SIZE AND COGNITIVE ABILITY

J. Philippe Rushton and C. Davison Ankney

INTRODUCTION

Enormous variability exists in brain size and cognitive ability within each of the populations to be discussed. Because group distributions overlap substantially on the variables in question, with average differences amounting to between 4 and 34 percent, it is problematic to generalize from group averages to individuals. Nor do we hold that brain-size differences are due entirely to genetic factors. Individual brain size (and cognitive ability) can obviously be affected by nutrition and early experience. We later describe a twin study of Whites and Blacks, boys and girls, estimating that only about 50 percent of cranial size variation is due to genetic factors (Rushton & Osborne, 1995). Nonetheless, the mean group differences we identify in this chapter seem to us of great interest.

Brain Size and Intelligence

Galton (1888) was one of the first to quantify the brain-size/cognitive-ability relationship within humans. Galton's subjects were 1,095 Cambridge

Advances in Cognition and Educational Practice, Volume 4, pages 25-77.

ISBN: 0-7623-0105-8

undergraduate men divided into those who had achieved first class honors degrees and those who had not. Galton computed head volume by multiplying head length by breadth by height and plotting the results against age (19 to 25 years) and class of degree (A, B, C). He reported that (1) cranial capacity continued to grow after age 19, and (2) men who obtained high honors degrees had a brain size from 2 to 5 percent greater than those who did not. Pearson (1906) re-examined Galton's data using his newly developed correlation coefficient and found a small positive relationship between head size and university grade. This has remained the general conclusion.

We (Rushton & Ankney, 1996) recently summarized 47 samples of the relation between brain size and cognitive ability (see Table 1). Clinical samples have been excluded except where identified in the section on imaging. The most representative or average correlation has been reported from those studies providing multiple correlations (e.g., by age and sex or by adjusting for body size). Double entries were eliminated, particularly those emanating from the U.S. National Collaborative Perinatal Project (Broman, Nichols, Shaughnessy, & Kennedy, 1987). Not included in Table 1 are typological studies showing that although gifted children typically have larger heads than average (Fisch, Bilek, Horrobin, & Chang, 1976; Terman, 1926/1959, p. 152), mentally defective children typically have smaller heads than average (Broman et al., 1987; Hack et al., 1991). Clinical exceptions to the general rule of larger brain/higher IQ includes autism and hydrocephalic retardation, perhaps because of the disease process (Piven, Arndt, Bailey, Havercamp, Andreasen, & Palmer, 1995).

The 47 samples are categorized into four sections. Section A shows results of 17 studies that took external head measurements from a total of 45,056 children and adolescents and correlated these with estimates of mental ability from ratings, grades, and standardized tests. Correlations ranged from 0.08 to 0.35 with an unweighted mean of 0.21 (when weighted by sample size, 0.20). Section B shows results from 15 studies of adult head-size/cognitive-ability (total $N = 6{,}437$ people). Correlations ranged from 0.02 to 0.39 with an unweighted mean of 0.15 (when weighted by sample size, also 0.15). Section C shows the results of 7 mainly clinical samples using a total of 312 adults with brain size estimated by Computer Assisted Tomography (CAT) and Magnetic Resonance Imaging (MRI) and cognitive ability estimated by educational achievement or by standardized tests. Correlations ranged from 0.07 to 0.38 with an unweighted mean of 0.24 (when weighted by sample size, 0.22). Section D shows the results of 8 non-clinical samples using a total of 381 adults with brain size estimated by CAT and MRI and cognitive ability estimated by educational and occupational achievement or by standardized tests. Correlations ranged from 0.33 to 0.69 with an unweighted mean of 0.44 (when weighted by sample size, 0.42). The exact p value of all 46 correlations was less than 10^{-10}.

Table 1. Head/Brain Size and Intelligence

Source	Sample	Measurement: Head/Brain	Measurement: Test	Correlation
A. Children and Adolescents by External Head Measurements				
Pearson (1906)	4,386 British Children (2,198 boys, 2,188 girls) aged 3 to 20; standardized to age 12	Length	Teachers' estimate	0.11
Murdock and Sullivan (1923)	595 American children aged 6 to 17 standardized by age and sex	Perimeter	IQ tests	0.22
Estabrooks (1928)	251 European-American children (102 boys, 149 girls) aged 6 years	Capacity	Binet	0.19
Porteus (1937)	200 White Australian children	Perimeter	Porteus Maze	0.20
Robinow (1968, cited in Jensen & Sinha, 1993)	300 boys and girls aged 3 to 13 studied longitudinally, with age standardized.	Perimeter	IQ tests	0.18
Klein et al. (1972)	170 Guatemalan Amerindian children aged 3 to 6	Perimeter	Knowledge test standardized with age-sex groups	0.28
Weinberg et al. (1974)	334 European-American boys aged 8 to 9 years	Perimeter	WISC	0.35
Swan et al. (undated, cited in Cattell, 1982, & Jensen & Sinha, 1993)	547 children, age controlled	Perimeter	IQ tests	0.11
Broman et al. (1987)	18,907 African-American boys and girls aged 7 years	Perimeter	WISC	0.19
	17,241 European-American boys and girls aged 7 years	Perimeter	WISC	0.24
Lynn (1990)	310 Irish boys and girls aged 9 to 10	Perimeter	PMAT	0.18

(continued)

Table 1. (*Continued*)

	205 Irish children aged 9 years	Perimeter	Matrices	0.26
	91 English children aged 9 years	Perimeter	Matrices	0.26
Osborne (1992)	224 European-American children (106 boys, 118 girls) aged 13 to 17; Controls for height and weight	Capacity	Basic	0.29
Lynn and Jindal (1993)	200 east Indian 9-year-olds from Kurukshetra in northern India (100 boys, 100 girls).	Perimeter	Matrices	0.15
Swan and Miszkiewicz (undated, cited in Jensen & Sinha, 1993).	843 European-American children in grades K-12, with age controlled	Perimeter	IQ tests	0.08
Summary of A	*Number of samples: 17* *Total N: 45,056* *Range of r: 0.08-0.35* *Unweighted mean r: 0.21* *N-weighted mean r: 0.20*			
B. Adults by External Head Measurements				
Pearson (1906)	1,011 British male university students	Length	Grades	0.11
Pearl (1906)	935 Bavarian male soldiers	Perimeter	Officers' ratings	0.14
Reid and Mulligan (1923)	449 Scottish male medical students	Capacity	Grades	0.08
Summerville (1924)	105 European-American male university students	Capacity	Thorndike	0.08
Wrzosek (1931, cited in Henneberg et al, 1985)	160 Polish male medical students	Capacity	Baley's Polish language IQ test	0.14

Schreider (1968)	80 Otomi Amerindians from Meixco of unreported sex	Perimeter	Form Board	0.39
	158 French peasants of unreported sex	Perimeter	Matrices	0.23
Passingham (1979)	415 English villagers (212 men, 203 women) aged 18 to 75	Capacity	WAIS	0.13
Susanne (1979)	2,071 Belgian male conscripts	Perimeter	Matrices	0.19
Henneberg et al. (1985)	302 Polish medical students (151 men, 151 women) aged 18 to 30 years.	Capacity	Bailey's Polish language IQ test	0.14
Bogaert and Rushton (1989)	216 White Canadian men and women university students, adjusted for sex	Perimeter	MAB	0.14
Rushton (1992b)	73 Asian Canadian men and women university students	Perimeter	MAB	0.14
	211 White Canadian men and women university students	Perimeter	MAB	0.21
Reed and Jensen (1993)	211 European-American men college students	Capacity	Various	0.02
Wickett et al. (1994)	40 White Canadian women university students	Perimeter	MAB	0.11
Summary of B	*Number of samples: 15* *Total N: 6,437* *Range of r: 0.02-0.39* *Unweighted mean r: 0.15* *N-weighted mean r: 0.15*			

(continued)

Table 1. (*Continued*)

C. Adult clinical samples by imaging techniques				
Yeo et al. (1987)	41 European-Americans (14 men, 27 women, mean age = 38) with medically unconfirmable neurologic symptoms	CAT brain area of 8 or 9 contiguous slices encompassig 53% of brain	WAIS	0.07
DeMyer et al. (1988)	45 schizophrenic patients and controls matched for age, race and sex but not education, with a mean age of 29 years	MRI, average of 4 slices	WAIS verbal	0.21
Andreasen et al. (1990)	54 mainly (99%) European-American schizophrenics (36 men, 18 women) with a mean age of 34 years	MRI frontal lobe	Educational level	0.31
Flaum et al. (1994)	72 schizophrenic patients (52 men and 22 women) with height controlled	MRI volume	WAIS-R	0.11
Harvey et al. (1994)	26 British bipolar patients (62% women, 65% caucasian) aged 21-49	MRI volume	NART	0.38
Harvey et al. (1994)	48 British schizophrenic patients (59% women, 65% Caucasian) aged 19-61	MRI volume	NART	0.24
Haier et al. (1995)	26 mixed mild mental retardation, Downs Syndrome, and college student controls (38% controls, 54% males) with a mean age of 28 years; Corrected for the extended IQ range	MRI volume	WAIS-R	0.36

Summary of C	*Number of sample: 7* *Total N: 312* *Range of r: 0.07-0.38* *Unweighted mean r: 0.24* *N-weighted mean r: 0.22*			
D. Adult non-clinical samples by imaging techniques				
Pearlson et al. (1989)	84 normal Americans of racially heterogeneous background (63% white, 35% men, with a mean age of 40) used as a control group for a study of schizophrenics	CAT area of a single slice	Occupational status	0.35[1]
Andreasen et al. (1990)	47 European-Americans (28 men, 19 women) with a mean age of 35 years	MRI frontal lobe	Educational level	0.33
Willerman et al. (1991)	40 European-American university students (20 men, 20 women); corrected for sex, body size and the extended IQ range	MRI volume	WAIS	0.35
Andreasen et al. (1993)	67 European-American adults (37 men, 30 women) with a mean age of 38	MRI volume	WAIS	0.38
Raz et al. (1993)	29 European-American adults (I17 men, 12 women) aged 18 to 78	MRI volume	CFIT	0.43
Harvey et al. (1994)	34 normal British controls (45% women, 62% Caucasian) aged 19 to 49 years	MRI volume	NART	0.69

(continued)

Table 1. *(Continued)*

Wickett et al. (1994)	40 white Canadian women aged 20 to 30 years; height and partialled out and corrected for restriction of range	MRI volume	MAB	0.54
Egan et al. (1995)	40 British military (unreported race and sex breakdown) aged 23 years; corrected for height, weight and restriicted range	MRI volume	WAIS-R	0.48
Summary of D	*Number of samples: 8* *Total N: 381* *Range of r: 0.33-0.69* *Unweighted mean r: 0.44* *N-weighted mean r: 0.42*			

Notes: CAT = Computer Assisted Tomography; CFIT = Culture Free Intellignece Test; MAB = Multidimensional Aptitude Battery; MRI = Magnetic Resonance Imaging; NART = New Adult Reading Test; PMAT = Primary Mental Abilities Test; WAIS-R = Wechsler Adult Intelligence Scale-Revised; WISC = Wechsler Intelligence Scale for Children.

[1] Eta correlation calculated from *F* ratio (Schultz, 1991).

Source: From Rushton and Ankney (1996). Copyright, 1996 by The Psychonomic Society. Reprinted with permission.

The U.S. National Collaborative Perinatal Project mentioned in Table 1 (Broman et al., 1987), is worth mentioning further because it showed that even head perimeter measured *at birth* predicted IQ scores *at age 7* in 17,000 White and 19,000 Black children with correlations of about $r = 0.13$. Further evidence that the relation between head-size and cognitive-ability is functional comes from two studies showing that the relation exists *within* families as well as *among* families. A tendency for a sibling with a larger head to have a higher IQ than a sibling with a smaller head is of special interest because it controls for many of the sources of variance that distinguish families such as cultural background and socioeconomic status. Jensen (1994) examined 82 pairs of monozygotic and 61 pairs of dizygotic adolescent twins and extracted the general factor, or psychometric *g*, from their IQ tests and found it correlated with head size across individuals ($r = 0.30$), within twin pairs ($r = 0.25$), and between twin pairs ($r = 0.32$). Jensen and Johnson (1994) examined the head-size/IQ relation in some 14,000 pairs of siblings, from the Collaborative Perinatal Project mentioned above. At age 7 (although not at age 4) the significant correlation existed within-families ($r = 0.11$) as well as between-families ($r = 0.20$).

It is reasonable to expect that brain size and cognitive ability are related because Haug (1987, p. 135) found a correlation of $r = 0.479$ ($n = 81, p < 0.001$) between number of cortical neurons (based on a partial count of representative areas of the brain) and brain size, including both men and women in the sample. The regression equating the two was given as (# of cortical neurons [in billions] = 5.583 + 0.006 [cm^3 brain volume]). This means that a person with a brain size of 1,400 cm^3 has, on average, 600 million fewer cortical neurons than an individual with a brain size of 1,500 cm^3. The difference between the low end of normal (1,000 cm^3) and the high end (1,700 cm^3) works out to be 4.283 billion neurons (a difference of 27% more neurons from a 41% increase in brain size). The human brain may contain up to 100 billion (10^{11}) nerve cells classifiable into 10,000 types resulting in 100,000 billion synapses (Kandel, 1991). Even storing information at the low average rate of one bit per synapse, which would require two levels of synaptic activity (high and low), the structure as a whole would generate 10^{14} bits. Contemporary supercomputers, by comparison, command a memory of about 10^9 bits of information.

Race Differences in Brain Size

From weighing wet brains at autopsy and calculating cranial capacity from skulls and external head measurements, we show, from modern as well as historical studies, that Mongoloids (East Asians and their descendents) and Caucasoids (Europeans and their descendents) average larger brains than Negroids (Africans and their descendents). The Mongoloid/Negroid finding is especially striking because Mongoloids are often small in body size. When

adjustments are made for body size, Mongoloids average even larger and heavier brains than do Caucasoids. Although sampling and methodological difficulties may be identified in particular studies, results obtained from multimethod comparisons allow a triangulation on probable truth.

Brain Weight at Autopsy

In an analysis highly critical of the early literature on wet brain mass measured at autopsy, Tobias (1970) claimed that all interracial comparisons were "invalid," "misleading," and "meaningless" because 14 crucial variables had been left uncontrolled. In one study or another, these included "sex, body size, age of death, nutritional state in early life, source of the sample, occupational group, cause of death, lapse of time after death, temperature after death, anatomical level of severance, presence or absence of cerebrospinal fluid, of meninges, and of blood vessels" (p. 3). Tobias pointed out that each of these variables alone could increase or decrease brain mass by 10 to 20 percent, an amount equivalent to or greater than any purported race difference. He equally opposed conclusions of race differences in structural variables such as cortical thickness, size of frontal lobe, or complexity of the brain's convolutions.

Rushton (1988a) countered that aggregating across studies typically cancels measurement error, at least non-systematic measurement error, and he (Rushton, 1988a) calculated the midpoints of the range of scores provided by Tobias (1970, p. 6, Table 2) and found that a "Mongoloid series" (Tobias's term) averaged 1,368 grams, Caucasoids 1,378 grams, and Negroids 1,316 grams. Rushton (1988b) also averaged a related measure that took body size into account, that is, the "millions of excess nerve cells" estimated by Tobias for 8 subgroups and nationalities (1970, p. 9, Table 3). These were the number of neurons available for general adaptive purposes over and above that necessary for maintaining bodily functioning and were derivable from equations based on brain/body-weight relationships (Jerison, 1963, 1973). Tobias was skeptical of the value of his "exercise" and provided few details. Nonetheless, Rushton (1988b) found, in millions of excess neurons, Mongoloids = 8,990, Caucasoids = 8,650, and Negroids = 8,550.

Subsequent to Tobias's (1970) review, a major autopsy study was carried out by Ho et al. (1980a, 1980b) who provided original brain weight data for 1,261 adult subjects aged 25 to 80 from Cleveland, Ohio, after excluding obviously damaged brains. Ho et al. avoided most of the problems cited by Tobias and reported that, for sex-combined samples, brain mass averaged 1,323 grams for 811 American Whites ($SD = 146$) and 1,223 grams for 450 American Blacks ($SD = 144$). Although Ho et al. (1980a) did not provide values corrected for age or body size, the race differences in absolute brain mass cannot be explained by those variables: Black men and women in the sample were, on average, virtually identical in age and size to their White counterparts.

In the introduction to their article, Ho et al. (1980a) briefly reviewed additional literature from which Rushton (1988a) calculated that Mongoloids averaged 1,334 grams, Caucasoids 1,307 grams, and Negroids, 1,289 grams. Averaging the three sets of estimates (Tobias's review, Ho et al.'s review, and Ho et al.'s data) Rushton (1988a) found a sex-combined brain-weight for Mongoloids of 1,351 grams, Caucasoids 1,336 grams, and Negroids 1,286 grams. Further, Ho et al.'s review suggested that whereas the Caucasoid brain weight began to decline at 25, the Mongoloid brain weight may not do so until age 35.

Endocranial Volume

Many more studies have estimated brain size from cranial capacity for, as Baker (1974, p. 429) remarked, "Skulls are many, freshly removed brains are few." The endocranial literature, however, also underwent serious critiques, as in Gould's (1978, 1981) reanalysis of Morton's (1849) data. In particular, Gould (1978, p. 503) alleged "unconscious...finagling" and "juggling" (1978, p. 503). In his widely cited *Mismeasure of Man*, which has become a standard reference guide to this literature, Gould (1981, p. 65) suggested how biases could be introduced into such data:

> Plausible scenarios are easy to construct. Morton, measuring by seed, picks up a threateningly large black skull, fills it lightly and gives it a few desultory shakes. Next, he takes a distressingly small Caucasian skull, shakes hard, and pushes mightily at the foramen magnum with his thumb. It is easily done, without conscious motivation; expectation is a powerful guide to action.

Table 2 represents Gould's summary of Morton's data after "correcting" Morton's alleged errors. The first column reports Gould's 1978 summary and the second column his 1981 summary following an admission by Gould (1981, p. 66) of his own "embarrassing" error in calculating the 1978 figures. In both his 1978 and 1981 writing, Gould dismissed the differences between groups as "trivial."

Rushton (1989a) averaged Gould's 1978 and his 1981 figures on cranial capacity and found, on both occasions, that Mongoloids (Native Americans + Mongoloids) > Caucasoids (Modern Caucasians + Ancient Caucasians) > Negroids (Africans). After excluding "Malays" due to uncertainty as to their racial category, the figures from column 1 are 85.5, 84.5, and 83 cubic inches (1,401, 1,385 and 1,360 cm^3) and from column 2 they are 86.5, 85.5, and 83 cubic inches, respectively (1,418, 1,401 1,360 cm^3). The figures did not change appreciably if Malays were included as either Mongoloids or Caucasoids. Clearly, despite Gould's conclusions, Mongoloids and modern Caucasians had an advantage of 4 cubic inches (64 cm^3) over Africans in these "corrected" data. Differences of even 1 cubic inch (16 cm^3) should probably not be dismissed

Table 2. S.J. Gould's "Corrected" Final Tabulation of Morton's Assessment of Racial Differences in Cranial Capacity

Population	*Cubic inches*	
	1978 Version	*1981 Version*
Native Americans	86	86
Mongolians	85	87
Modern Caucasians	85	87
Malays	85	85
Ancient Caucasians	84	84
Africans	83	83

Source: From Rushton (1989a, p. 14, Table 2). Copyright 1989 by Academic Press. Reprinted with permission.

as "trivial." As we have reviewed, Haug's (1987) calculations show that a 16 cm^3 difference translates into millions of neurons and hundreds of millions of synapses.

In any case, Gould's charge that Morton "unconsciously" doctored his results to show Caucasian racial superiority has been called into question. A random sample of the Morton collection was remeasured by Michael (1988) who found that very few errors had been made and that these were not in the direction that Gould had asserted. Instead, errors were found in Gould's own work. Michael (1988, p. 353) concluded that Morton's research "was conducted with integrity... (while)...Gould is mistaken."

Rushton (1988a) also averaged other data on endocranial volume and found support for his ranking based on Gould's analyses. For example, Coon (1982) had calculated cranial capacities for 17 populations from detailed measurements made by Howells (1973) of 2,000 skulls recorded on a tour of the world's museums. Coon had concluded that "Asiatic Mongols, Eskimoes, and Polynesians have the largest brains, European Caucasoids the next largest, Africans and Australoids still smaller, and the small or dwarfed peoples the smallest" (1982, p. 18). Coon's book began with a Preface from Howells warning readers not to be too easily dismissive. Combining the sexes, Rushton (1988a) found that Mongoloids = 1,401 cm^3, Caucasoids = 1,381 cm^3, and Negroids = 1,321 cm^3. Rushton (1988a) also averaged capacities from a table provided by Molnar (1983, p. 65) based on data from Montagu (1960) and found that Mongoloids = 1,494 cm^3, Caucasoids = 1,435 cm^3, and Negroids = 1,346 cm^3. Rushton (1988a) then averaged across Coon's and Molnar's figures to find Mongoloids = 1,448 cm^3, Caucasoids = 1,408 cm^3, and Negroids = 1,334 cm^3.

Analyzing the world database of about 20,000 skulls, uncorrected for body size, Beals, Smith and Dodd (1984, p. 307, Table 5) found that the size of the sex-combined brain cases differed by continental area. Overall there was a 2.5 cm^3 increase in brain volume with each degree of latitude. Regional

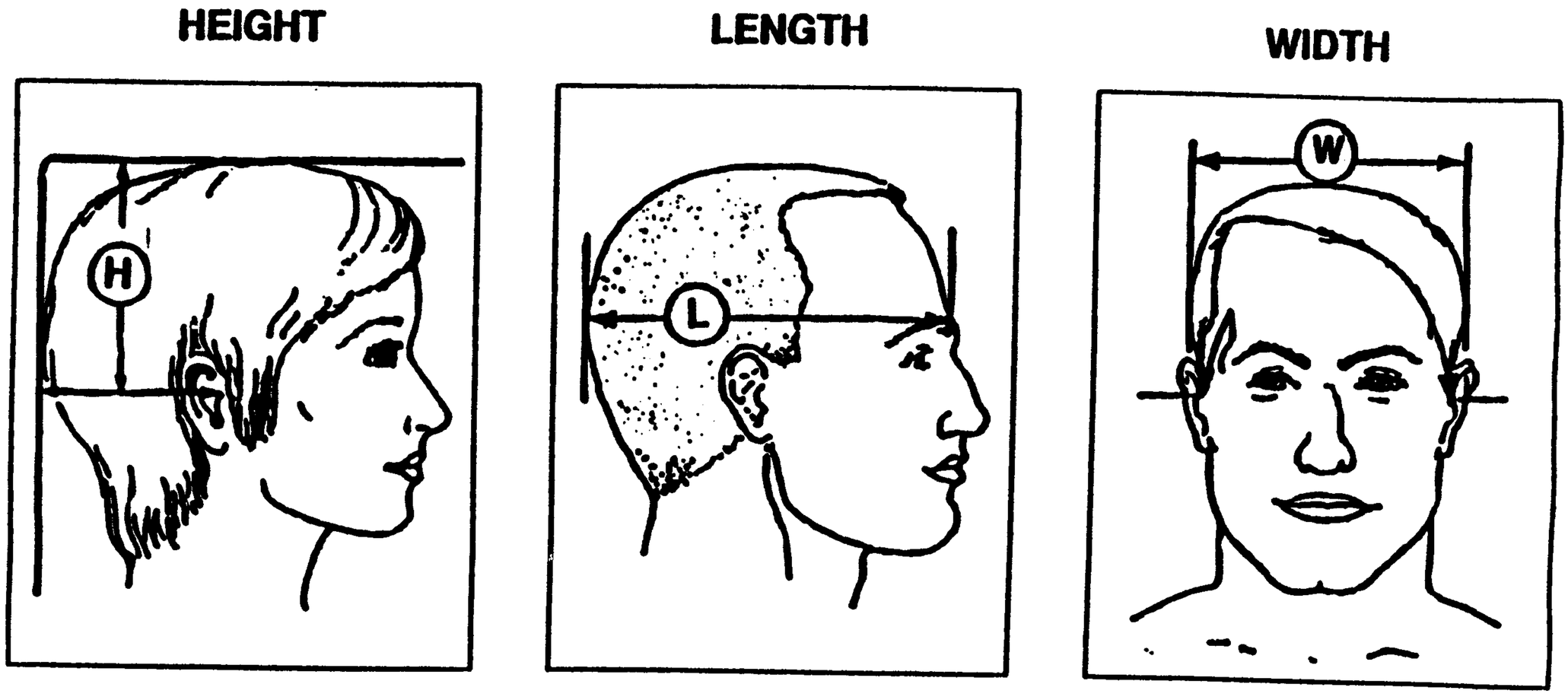
differences emerged. Excluding Caucasoid areas of Asia (e.g., India) and Africa (e.g., Egypt), 19 Asian populations averaged 1,415 cm^3 ($SD = 51$), 10 European groups averaged 1,362 cm^3 ($SD = 35$), and 9 African groups averaged 1,268 cm^3 ($SD = 85$).

External Head Measurements

A third way of estimating cranial capacity is from external head measurements (Figure 1). For example, the length, width and height of the head are placed in regression equations to predict cranial capacity. Lee and Pearson (1901) may have been the first to do this.

They chose skulls in series of 50 to 100 from widely different races to permit generalizing the results. Capacities had been determined independently by competent observers. Altogether, skull dimensions based on the greatest length, the greatest breadth, and the height measured from the auricular line were examined for 941 men and 516 women including representatives of the Asian, European and African continents (p. 246, Table XX).

Lee and Pearson (1901) showed that their equations, including a "panracial" equation (p. 252, Number 14, p. 260), provided cranial capacity estimates more accurate for the individual skull than the direct method of using sand, seed, or shot. Their equations predicted both male and female capacities with errors of less than 1 percent, or about 2 to 5 cm^3 on crania of 1,300 to 1,500 cm^3 (p. 244, Table XVIII), considerably less than the 30 cm^3 difference typically found between two observers measuring the same series of skulls using an internal "packing" procedure.



Notes: *H = height, L =* length, and *W =* width.
Source: From *Race, Evolution, and Behavior* (p. 117, Figure 6.1), by J.P. Rushton, 1995, New Brunswick, NJ: Transaction. Copyright 1995 by Transaction Publishers. Reprinted with permission

Figure 1. Cranial Capacity Estimated from External Head Measurements

Lee and Pearson's (1901) panracial equations were:
for men,

$$CC\,(\text{cm}^3) = 0.000337\,(L\text{-}11\text{mm})(B\text{-}11\text{mm})(H\text{-}11\text{mm}) + 406.01 \qquad (1)$$

for women,

$$CC\,(\text{cm}^3) = 0.0004\,(L\text{-}11\text{mm})(B\text{-}11\text{mm})(H\text{-}11\text{mm}) + 206.6 \qquad (2)$$

where CC is cranial capacity and L, B, and H are length, breadth and height in millimeters and 11mm is subtracted for fat and skin around the skull. When data for height of head are missing, cranial capacity can be estimated from another equation given by Lee and Pearson (1901, p. 235, Table VII, Number 5; as used by Passingham (1979) and amended by Rushton (1993) to subtract 11mm for fat and skin around the skull):

for men,

$$CC\,(\text{cm}^3) = 6.752\,(L\text{-}11\text{mm}) + 11.421\,(B\text{-}11\text{mm}) - 1{,}434.06 \qquad (3)$$

for women,

$$CC\,(\text{cm}^3) = 7.884\,(L\text{-}11\text{mm}) + 10.842\,(B\text{-}11\text{mm}) - 1{,}593.96 \qquad (4)$$

These equations have been applied to several sets of anthropometric data. One set, compiled by Herskovits (1930), has often been cited as showing an absence of racial differences because of the amount of overlap in the distributions. As can be seen in Table 3, a sample of 961 American Negroes had larger head measurements than a sample of Swedes. However Herskovits's monograph actually contained information on head length and width for 26 internatioanl samples ($N = 54{,}454$; males only). Using equation 3, Rushton (1993) calculated cranial capacities for each sample and found a highly significant difference in the predicted direction. Five "Mongoloid" samples (in this case, mostly North American Indians) averaged 1,451 cm^3 ($SD = 22$), 9 Caucasoid samples averaged 1,421 cm^3 ($SD = 49$), and 12 Negroid samples averaged 1,295 cm^3 ($SD = 44$). No information was available on body size.

Herskovits's (1930) data were gathered by different investigators from different parts of the world using different techniques. Although not too much reliance should be placed on the reanalysis of this one study, nonetheless, because Herskovits's monograph has so often been referred to by those critical of race differences, the aggregation is noteworthy. It strikingly confirms the re-aggregations of the "corrected" data sets by Tobias (1970) and Gould (1978, 1981).

Table 3. Cranial Capacities Calculated from Head Lengths and Widths Provided by Herskovits (1993) For Various Male Samples and Classified by Race or Geographical Region

		Length (mm)	*Width (mm)*	*Capacity (cm^3)*
Mongoloids and Asians				
540	Pure Sioux	194.90	155.10	1,453
77	Half-blood Sioux	194.40	154.30	1,441
50	Montagnais-Naskapi	194.00	157.10	1,470
83	Marquesans	193.20	153.20	1,420
86	Hawaiians	191.25	158.93	1,472
	Mean	193.55	155.73	1,451
Caucasoids and Europeans				
727	Old Americans	197.28	153.76	1,454
263	Scotch foreign-born	196.70	153.80	1,451
959	Oxford students	196.05	152.84	1,435
493	Aberdeen students	194.80	153.40	1,433
46,975	Swedes	193.84	150.40	1,393
1,000	Cambridge students	193.51	153.96	1,431
802	Cairo natives	190.52	144.45	1,302
450	Foreign-born Bohemians	189.80	159.10	1,465
60	American-born Bohemians	188.00	156.50	1,423
	Mean	193.39	153.13	1,421
Negroids and Africans				
961	American Negroes	196.52	151.38	1,422
91	Masai	194.67	142.49	1,308
34	Lotuko	192.90	141.30	1,283
55	Kajiji	192.31	144.56	1,316
27	Somali	191.81	143.19	1,297
19	Ekoi	191.05	143.16	1,297
40	Vai	188.85	142.45	1,268
384	Akikuyu	188.72	143.25	1,276
72	Kagoro	188.19	142.43	1,263
128	Akamba	187.80	143.63	1,275
48	Ashanti	187.33	145.01	1,287
30	Acholi	187.30	141.80	1,250
	Mean	190.62	143.72	1,295

Notes: Cranial capacities (cm^3) = 6.752 x (*L - 11mm*) + 11.421 x (*W* - 11mm) - 1434.06. Formula is from Lee and Pearson (1901).

Source: From Rushton (1993, p. 230, Table 1). Copyright 1993 by Pergamon Press. Reprinted with permission.

The body size and head size data for 24 international male military samples totalling 57,378 individuals were abstracted by Rushton (1991) from a compilation made by the U.S. National Aeronautics and Space Administration. For each sample, cranial capacities were calculated using equation 3. The unadjusted cranial capacity for the 4 Mongoloid samples

Table 4. Cranial Capacity (cm^3) by Sex, Rank, and Race for 6,235 U.S. Military Personnel: Range of Least Square Means and Midpoint of Range from all Analyses Wherein Race, Sex, and Rank were Examined Together

Sex and Rank, Race	*Range (All Groups - Corrected for Body Size)*	*Midpoint of Range*
Female, Enlisted		
Negroid	1,288 - 1,319	1,304
Caucasoid	1,299 - 1,330	1,315
Mongoloid	1,326 - 1,387	1,357
Female, Officer		
Negroid	1,293 - 1,324	1,309
Caucasoid	1,315 - 1,346	1,331
Mongoloid	1,341 - 1,417	1,379
Male, Enlisted		
Negroid	1,405 - 1,424	1,415
Caucasoid	1,419 - 1,445	1,432
Mongoloid	1,464 - 1,473	1,469
Male, Officer		
Negroid	1,400 - 1,438	1,419
Caucasoid	1,429 - 1,453	1,441
Mongoloid	1,463 - 1,487	1,475

Source: From Rushton (1992a, p. 408, Table 3). Copyright 1992 by Ablex Publishing Corporation. Reprinted with permission.

was 1,343 cm^3 ($SD = 47$) and for 20 Caucasoid samples, 1,467 cm^3 ($SD = 58$). The stature, weight, and total body surface area of the Mongoloid samples were all significantly lower than those of the Caucasoid samples. After adjusting for the body size variables, the least-square mean for Mongoloids was 1,460 cm^3 and for Caucasoids 1,446 cm^3, although the comparison was not statistically significant.

Analyses reaching high levels of statistical significance were conducted on a stratified random sample of 6,325 U.S. Army personnel (Rushton, 1992a). Individual and head measurements were available separately for men and women, officers and enlisted personnel, and those who had defined themselves to the U.S. Army as Black, Asian or White. Because head measurements were available for length, width and height, cranial capacities were calculated from Equations (1) and (2). For the entire sample, the unadjusted size of the cranium was 1,375 cm^3. The range was from 981 cm^3, a Black woman, to 1,795 cm^3, a White man. Because the measurements had been gathered on individuals, specific adjustments could be made to the raw data for the effects of age, stature and weight, and then sex, rank or race.

Table 5. Cranial Capacities of World Populations of 25-45 Years-Olds

Region, number of references, and countries	Men				Women			
	Stature (mm)	Head length (mm)	Head breadth (mm)	Cranial capacity[a] (cm^3)	Stature (mm)	Head length (mm)	Head breadth (mm)	Cranial capacity[b] (cm^3)
1. North America (34 refs. from Canada and United States)	1,790	195	155	1,453	1,650	180	145	1,191
2. Latin America (20 refs. from Indian populations from Bolivia, Honduras, Peru, etc).	1,620	185	150	1,328	1,480	175	145	1,152
3. Latin America (15 refs. from European-Negroid populations in Argentina, Chile, the Caribbean Island States, etc).	1,750	190	155	1,419	1,620	175	150	1,206
4. Northern Europe (28 refs. from Denmark, Germany, Sweden, etc.)	1,810	195	155	1,453	1,690	180	150	1,246
5. Central Europe (42 refs. from Austria, Switzerland, etc.).	1,770	190	155	1,419	1,660	180	145	1,191
6. Eastern Europe (14 refs. from Poland and Soviet Union).	1,750	190	155	1,419	1,630	180	150	1,248

(continued)

Table 5. (Continued)

7.	South-Eastern Europe (40 refs. from Bulgaria, Greece, Romania, etc.).	1,730	190	155	1,419	1,620	175	150	1,206
8.	France (20 refs.)	1,770	195	155	1,453	1,630	180	140	1,137
9.	Iberian Peninsula (6 refs. from Spain and Portugal).	1,710	185	155	1,385	1,600	180	150	1,246
10.	North Africa (10 refs. from Algeria, Ethiopia, Niger, Sudan, etc.).	1,690	190	145	1,305	1,610	185	140	1,177
11.	West Africa (10 refs. from Congo, Ghana, Nigeria, Zaire, etc.).	1,670	195	145	1,339	1,530	180	135	1,083
12.	South-Eastern Africa (16 refs. from Angola, Kenya, Uganda, Zambia, etc.).	1,680	195	145	1,339	1,570	180	135	1,083
13.	Near East (5 refs. from Iraq, Lebanon, Turkey, etc.).	1,710	190	150	1,362	1,610	180	140	1,137
14.	North India (23 refs from Bangladesh, Nepal, Pakistan, etc.).	1,670	190	145	1,305	1,540	180	135	1,083
15.	South India (23 refs. from India, Maldives and Sri Lanka).	1,620	180	145	1,237	1,500	175	130	989
16.	North Asia (5 refs. from China, Mongolia, etc.).	1,690	190	150	1,362	1,590	180	145	1,191

17.	South China (9 refs. from Hong Kong, Macao, Taiwan, etc.).	1,660	190	150	1,362	1,520	180	145	1,191
18.	South-East Asia (11 refs. from Brunei, Indonesia, Malaysia, Philippines, etc.).	1,630	185	145	1,271	1,530	175	135	1,043
19.	Australia (6 refs. from European population in Australia and New Zealand).	1,770	192	155	1,433	1,670	180	145	1,191
20.	Japan (26 refs. from Japan and Korea).	1,720	190	155	1,419	1,590	180	145	1,191

Notes: [a]Cranial capacity for men (cm^3) = 6.752 (Head length - 11 *mm*) + 11.421 (Head Breadth - 11*mm*) - 1434.06
[b]Cranial capacity for women (cm^3) = 7.884 (Head length - 11 *mm*) + 10.842 (Head Breadth - 11*mm*) - 1593.96

Source: From Rushton (1994, pp. 286-287, Table 1).

The races differed significantly in both the unadjusted (raw) and adjusted cranial capacities. Analysis of variance of unadjusted cranial capacity showed that 543 Asian Americans averaged 1,391 cm^3 ($SD = 104$), 2,871 European Americans, 1,378 cm^3 ($SD = 92$) and 1,387 African Americans, 1,362 cm^3 ($SD = 95$). After adjusting for the effects of stature, weight, sex, and military rank, the differences became larger with Asian Americans averaging 1,416 cm^3, European Americans 1,380 cm^3 and African Americans 1,359 cm^3. Attempts to diminish the differences in cranial capacity by numerous corrections for body size were unsuccessful. The mean values change with particular adjustments for body size, rank or sex (see Table 4). But in no case was there departure from the significant racial ranking.

Another study using external head measurements was carried out by Rushton (1994) based on a 1990 review of ergonomically important body measurements compiled by the International Labour Office in Geneva (Jurgens, Aune, & Pieper, 1990). Head and body size measurements had been gathered over a 30-year period for tens of thousands of men and women aged 25 to 45 years. Some 300 references had been examined from 7 sources: handicraft workers such as tailors and shoemakers, anthropology studies, medical records, sports participation, growth surveys, forensic and legal investigations, and ergonomic studies. The data had been grouped into the 20 world regions summarized in Table 5 which present the 50th percentile measures of stature, head length, and head breadth, separately for men and women as well as the number of references cited to arrive at the summarized figures. From these Rushton (1994) derived cranial capacities using equations (4) and (5) above.

Because the regions were fully described in the research report with respect to the included countries (see Table 5), it was possible to eliminate ambiguous categories thereby facilitating racial comparisons. Excluded from statistical analysis were 6 Regions: No. 2 (Amerindians), Nos. 3 and 10 (combining Caucasoid and Negroid populations), Nos. 14 and 15 (North and South India), and No. 18 (combining Caucasoid and Mongoloid populations). With each of the male/female sample means treated as an independent entry, there are 6 clear Mongoloid samples (Regions 16, 17, and 20), 18 predominately Caucasoid samples (Regions 1, 4, 5, 6, 7, 8, 9, 13 and 19), and 4 clear African samples (Regions 11 and 12). Analysis of variance carried out on the unadjusted (raw) means showed that the East Asians ($M = 1{,}286$ cm^3, $SD = 117$) and Caucasoids ($M = 1{,}311$ cm^3, $SD = 103$) averaged larger absolute cranial capacities than did Africans ($M = 1{,}211$ cm^3, $SD = 144$). After adjusting for the effects of stature, the differences became more pronounced with East Asians averaging 1,308 cm^3; Caucasians, 1,297 cm^3; and Africans 1,241 cm^3. Subsidiary analyses weighted by the number of references or some other combination of countries did not alter the overall pattern of the results.

Summary of Brain Size Data

Table 6 (from Rushton, 1995) summarizes the results from 44 studies of race differences in adult brain size. Four sets of data are itemized and averaged. Section A sets out the results of brain mass (grams) from autopsy studies. There were 38 of these, including 16 reports of data from Mongoloids in Korea and Japan, 18 from Caucasoids in Europe and the United States, and 8 from Negroids in Africa and the United States. The results are shown for men and women separately where possible. For some studies, key reviews were used because the originals were unpublished, in a foreign language, or otherwise difficult to obtain. In the studies cited by Dekaban and Sadowsky (1978), Rushton (1995) calculated the mean as the midpoint of a range. Double entries were eliminated and all measures converted to cm^3 using an equation from Hofman (1991). After the 38th autopsy study, summary statistics are provided for each racial group showing the number of studies, the range, the median and the mean. The sex combined averages are calculated by adding the mean and median figures for men and women and dividing by two.

The results in section A show that the sex-combined mean brain weight of Mongoloids is almost as heavy (1,304 g) as that of Caucasoids (1,309 g) and that both of these are higher than those of Negroids (1,180 g). The statistical significance of these difference can be guaged from the fact that in no case is the brain weight of a Negroid sample of men or women above the mean or median of those of Mongoloids or Caucasoids ($p < .001$). Translating the grams into cm^3, the Mongoloids, Caucasoids and Negroids average, respectively: 1,351, 1,356, 1,223 cm^3.

Section B sets out the endocranial data. Here the global review by Beals et al. (1984) based on up to 20,000 endocranial specimens from 122 ethnic groups is relied on. The sex-combined mean for Mongoloids is 1,415 cm^3, for Caucasoids, 1,362 cm^3, and for Negroids, 1,268 cm^3. (Figures from Table 5 of Beals et al., 1984; sex differences from K. Beals, personal communication to J.P. Rushton, May 9, 1993). We did not include several endocranial studies carried out within the United States (e.g., Todd, 1923; Simmons, 1942), although the results are directly in line with those of Beals et al. (1984). Nor did we include a confirmation of Beals et al. from the independent review by Ricklan and Tobias (1986), who found that 917 Negroid males averaged 1,365 cm^3 and 320 Negroid females averaged 1,218 cm^3 for a sex-combined mean of 1,292 cm^3. Because of the degree of possible overlap in some series, we took Beals et al. (1984) to be sufficient.

Section C sets out four studies estimating cranial capacity from external head measurements. The non-asterisked figures are uncorrected for body-size while the asterisked figures have been corrected. As in Section A, the number of studies, the range, the mean and the median are provided. The uncorrected sex-combined mean cranial capacity of Mongoloids (1,335 cm^3) is virtually the

Table 6. Summary of Race Differences in Brain Size: Multimethod Comparisons

Source	Samples and Procedures	Mongoloids			Caucasoids			Negroids		
		Men	Women	Combined	Men	Women	Combined	Men	Women	Combined
A. *Autopsy Data (grams)*										
1. Peacock (1865, cited and averaged by Pearl, 1934)	5 Negro men.	—	—	—	—	—	—	1,257	—	—
2. Russell (1869, analyzed by Pearl, 1934)	379 Black soldiers and 24 White soldiers most of whom died of pneumonia during U.S. Civil War.	—	—	—	1,471	—	—	1,342	—	—
3. Doenitz (1874, cited in Spitzka, 1903)	10 Japanese men executed by decapitation.	1,337	—	—	—	—	—	—	—	—
4. Bischoff (1880, cited in Pakkenberg & Voigt, 1964)	906 Europeans measured at pathology institute.	—	—	—	1,362	1,219	(1,291)	—	—	—
5. Taguchi (1881, cited in Spitzka, 1903)	100 Japanese men executed by decapitation.	1,356	—	—	—	—	—	—	—	—
6. Topinard (1885, cited in Peal, 1934)	29 unspecified Negro men, collected from the literature.	—	—	—	—	—	—	1,234	—	—
7. Suzuki (1892, cited in Shibata, 1936)	27 Japanese (24 men, 3 women) aged 35 to 73.	1,348	1,120	(1,234)	—	—	—	—	—	—
8. Taguchi (1892, cited in Shibata, 1936)	524 Japanese (374 men, 150 women) aged 21 to 95.	1,367	1,214	(1,291)	—	—	—	—	—	—
9. Marshall (1892)	2,012 British (972 men, 1,040 women) aged 20 to 90 years; reanalysis of 1861, data published with breakdowns for age, height, weight, sane/insane.	—	—	—	1,329	1,194	(1,262)	—	—	—
10. Waldeyer (1894, cited in Pearl, 1934)	12 African men aged 15 or over.	—	—	—	—	—	—	1,148	—	—

(continued)

Table 6. (*Continued*)

11.	Retzius (1900, cited in Pakkenberg & Voigt, 1964)	700 Swedes at a pathological institute.	—	—	—	1,399	1,248	(1,324)	—	—	—
12.	Matiegka (1902, cited in Spakkenberg & Voigt, 1964)	416 Europeans at a pathological institute.	—	—	—	1,347	1,204	(1,276)	—	—	—
13.	Matiegka (1902, cited in Pakkenberg & Voigt, 1964)	581 Europeans at an Institute of Forensic Medicine in Prague.	—	—	—	1,450	1,306	(1,378)	—	—	—
14.	Marchand (1902, cited in Pakkenberg & Voigt, 1964)	1,169 Europeans aged 18 to 50 at a pathological institute in Marburg.	—	—	—	1,400	1,275	(1,338)	—	—	—
15.	Spitzka (1903)	597 Japanese (421 men, 176 women) aged 21 to 95 from hospitals around Tokyo; ten years of records including data on age, stature and weight.	1,367	1,214	(1,291)	—	—	—	—	—	—
16.	Bean (1906)	Review of records for 22 Negro men and 10 Negro women.	—	—	—	—	—	—	1,256	980	(1,118)
17.	Bean (1906)	125 Americans from an anatomical laboratory in Baltimore (37 White men, 9 White women, 51 Black men, 28 Black women).	—	—	—	1,341	1,103	(1,222)	1,292	1,108	(1,200)
18.	Chernyshev (1911, cited in Dekaban & Sadowsky, 1978)	Unspecified number of men and women (probably Russian) aged 20 to 80 years.	—	—	—	1,346	1,210	(1,278)	—	—	—
19.	Nagayo (1919, 1925, cited in Shibata, 1936)	485 Japanese (329 men, 156 women) aged 16 to 60.	1,362	1,242	(1,302)	—	—	—	—	—	—

20.	Kurokawa (1920, cited in Shibata, 1936)	440 Japanese (240 men, 200 women) aged 15 to 50.	1,402	1,256	(1,329)	—	—	—	—	—	—
21.	Kubo (1922, cited in Shibata, 1936)	60 Koreans (56 men, 4 women) aged 21 to 74.	1,353	1,206	(1,280)	—	—	—	—	—	—
22.	Kimura (1925, cited in Shibata, 1936)	405 Japanese (243 men, 162 women) aged 15 to 50.	1,402	1,249	(1,326)	—	—	—	—	—	—
23.	Muhlmann (1927, cited in Dekaban & Sadowsky, 1978)	Unspecified number of men and women (probably German) aged 20 to 80 years.	—	—	—	1,346	1,205	(1,276)	—	—	—
24.	Yoshizawa (1929, 1930, cited in Shibata, 1936)	315 Japanese (211 men, 104 women) aged 16 to 80.	1,361	1,231	(1,296)	—	—	—	—	—	—
25.	Hoshi (1930, cited in Shibata, 1936)	954 Japanese (551 men, 403 women) aged 16 or over.	1,396	1,255	(1,326)	—	—	—	—	—	—
26.	Hoshi (1930, cited in Shibata, 1936)	Unknown number of Japanese of both sexes aged 15 to 50.	1,406	1,261	(1,334)	—	—	—	—	—	—
27.	Amano-Hayashi (1933, cited in Shibata, 1936)	1,817 Japanese (1,074 men, 743 women) aged 16 or over.	1,375	1,244	(1.310)	—	—	—	—	—	
28.	Kusumoto (1934, cited in Shibata, 1936)	522 Japanese (342 men, 180 women) of unknown age.	1,360	1,241	(1,301)	—	—	—	—	—	—
29.	Vint (1934)	389 adult Kenyans of Bantu and Nilotic stock autopsied by author from native hospitals in Nairobi; only brains judged normal used; weights validated against cranial capacity using water technique and compared to data published on Europeans.	—	—	—	1,428	—	—	1,276	—	—

(*continued*)

Table 6. (Continued)

Source	Samples and Procedures	Mongoloids			Caucasoids			Negroids		
		Men	Women	Combined	Men	Women	Combined	Men	Women	Combined
30. Shibata (1936)	153 Koreans (136 men, 17 women) aged 15 to 78 years; those who died of diseases know to influence brain weight were excluded.	1,370	1,277	(1,324)	—	—	—	—	—	—
31. Roessle and Roulet (1938, cited in Pakkenberg and Voigt, 1964)	456 German soldiers	—	—	—	1,405	—	—	—	—	—
32. Appel and Appel (1942)	2,080 White U.S. men aged 12 to 96 at a mental hospital in Washington, DC; weights recorded from hospital records; brains with lesions and abnormalities discarded.	—	—	—	1,305	—	—	—	—	—
33. Takahashi and Suzuki (1961)	470 Japanese (301 men, 169 women) aged 30 to 69.	1,397	1,229	(1,313)	—	—	—	—	—	—
34. Pakkenberg and Voig (1964)	1,026 Danes (724 me, 302 women) aged 19 to 95 at the Forensic Institute in Copenhagen between 1959 and 1962; age, height, weight and cause of death examined.	—	—	—	1,440	1,282	(1,361)	—	—	—
35. Spann and Dustmann (1965, cited in Dekaban and Sadowsky, 1978)	Unspecified number of German men and women aged 15 to 94.	—	—	—	1,403	1,268	(1,336)	—	—	—
36. Chrzanowska and Beben (1973, cited in Dekaban and Sadowsky, 1978)	1,670 Poles (896 men, 774 women) aged 20 to 89.	—	—	—	1,413	1,266	(1,340)	—	—	—

37.	Dekaban and Sadowsky (1978)	4,736 U.S. Whites (2,733 males, 1,963 females) from hospitals around Washington, DC, aged birth to 86+; figures calculated for 16 to 86 years (2,036 men, 1,411 women).	—	—	—	1,392	1,254	(1,323)	—	—	—
38.	Ho et al. (1980a, 1980b)	1,261 White and Black Americans aged 25 to 80 (416 White men, 228 Black men, 395 White women, 222 Black women); weights taken from 5 years of records at Case Western Reserve University.	—	—	—	1,392	1,252	(1,322)	1,286	1,158	(1,222)
	Summary of A.	Number of studies	16	14	14	18	14	14	8	3	3
		Range	1,337-1,406	1,120-1,277	1,234-1,334	1,305-1,471	1,103-1,306	1,222-1,378	1,148-1,342	980-1,158	1,118-1,222
		Mean in grams	1,372	1,231	1,304	1,387	1,235	1,309	1,261	1,082	1,180
		Median in grams	1,367	1,242	1,306	1,396	1,250	1,323	1,267	1,108	1,200
		Mean in cm^3	1,421	1,275	1,351	1,437	1,280	1,356	1,306	1,121	1,223

(continued)

Table 6. (Continued)

Source	Samples and Procedures	Mongoloids			Caucasoids			Negroids		
		Men	Women	Combined	Men	Women	Combined	Men	Women	Combined
B. Endocranial Volume (cm^3)										
39. Beals et al. (1984)	Sex-combined endocranial volume from 122 populations based on up to 20,000 specimens from around the world and their geographic and climatic coordinates; packing was made with mustard seed; a standard 6% reduction made for studies reporting results based on lead shot.	1,491	1,340	(1,415)	1,441	1,283	(1,362)	1,338	1,191	(1,268)
C. Cranium Size from External Head Measurements (cm^3)										
40. Rushton (1990, amended 1993)	26 male populations (5 "Mongoloid"—mostly Ameridian, 9 European and European-American and 12 African and African-American; 54, 454 individuals) measurements compiled by Herskovits (1930).	1,451	—	—	1,421	—	—	1,295	—	—
41. Rushton (1991)	24 international male military samples (4 Mongoloid, 20 Caucasoid; 57,378 individuals). Head measurements compile by NASA (United States, 1978).	1,343 1,460*	—	—	1,467 1,446*	—	—	—	—	—
42. Rushton (1992a)	6,325 U.S. military personnel from a stratified random sample including officers and enlisted personnel, (411 Asian men, 132 Asian women, 1,590 White men, 1,281 White women, 1,381 Black men, 1,295 Black women); measurements gathered by the Army. (*Adjustments made for body size).	1,465 1,486*	1,300 1,319*	(1,383) (1,403)*	1,473 1,462*	1,268 1,259*	(1,371) (1,361)*	1,450 1,441*	1,261 1,250*	(1,356) (1,346)*

Table 6. (*Continued*)

43. Rushton (1994)	28 world samples (3 of Asian men, 3 of Asian women, 9 of Caucasian men, 9 of Caucasian women, 2 of African men, 2 of African women, tens of thousands of individuals); measurements compiled by the international Labour Office in Geneva. (*Adjustments made for body size).	1,381 1,371*	1,191 1,244*	(1,286) (1,308)*	1,422 1,378*	1,199 1,215*	(1,311) (1,297)*	1,339 1,337*	1,083 1,144*	(1,211) (1,241)*
Summary of C.	Number of studies (uncorrected)	4	2	2	4	2	2	3	2	2
	Range	1,343-1,465	1,191-1,300	1,286-1,383	1,421-1,473	1,199-1,268	1,311-1,371	1,295-1,450	1,083-1,261	1,211-1,356
	Mean in cm^3	1,410	1,246	1,335	1,446	1,234	1,341	1,361	1,172	1,284
	Median in cm^3	1,416	1,246	1,335	1,445	1,234	1,341	1,339	1,172	1,284
	Number of studies (corrected)	3	2	2	3	2	2	2	2	2
	Range	1,371-1,486	1,244-1,319	1,308-1,403	1,378-1,462	1,215-1,259	1,297-1,361	1,337-1,441	1,144-1,250	1,241-1,346
	Mean in cm^3	1,439	1,282	1,356	1,425	1,237	1,329	1,389	1,197	1,294
	Median in cm^3	1,460	1,282	1,356	1,446	1,237	1,329	1,389	1,197	1,294
D. Grand Summary: Mean of Means (cm^3)										
	Autopsies	1,421	1,275	1,351	1,437	1,280	1,356	1,306	1,121	1,223
	Endocranial volume	1,491	1,340	1,415	1,446	1,283	1,362	1,338	1,191	1,268
	External head measures	1,410	1,246	1,335	1,446	1,234	1,341	1,361	1,172	1,284
	Corrected external head measures	1,439	1,282	1,356	1,425	1,237	1,329	1,389	1,197	1,294
	Grand Mean	**1,440**	**1,286**	**1,364**	**1,437**	**1,259**	**1,347**	**1,349**	**1,170**	**1,267**

Note: *Adjustments for body size.

Source: From *Race, Evolution, and Behavior* (pp. 127-130, Table 6.6) by J.P. Rushton, 1995, New Brunswick, NJ: Transaction. Copyright 1995 Reprinted with permission.

same as for Caucasoids (1,341 cm^3), both of which average larger than Negroids (1,284 cm^3). Using the body size corrected figures in Section C shows Mongoloids average 1,356 cm^3, Caucasoids 1,329 cm^3, and Negroids 1,294 cm^3. These differences are significant within studies.

No exact solution is possible, of course, to the problem of how large the racial differences are in brain size. There is much variability from sample to sample with a clear overlap of distributions. Nonetheless, the consistency of results found using different procedures is noteworthy. In cm^3 the data from (a) autopsies, (b) endocranial volume, (c) head measurements, and (d) head measurements corrected for body size show: Mongoloids = 1,351, 1,415, 1,335, 1,356 (Mean = 1,364); Caucasoids = 1,356, 1,362, 1,341, 1,329 (Mean = 1,347); and Negroids = 1,223, 1,268, 1,284, 1,294 (Mean = 1,267). From these a world average brain size can be calculated of 1,326 cm^3, comparable to one of 1,349 cm^3 computed by Beals et al. (1984). The overall mean Asian/European difference favoring Asians was 17 cm^3, and the overall mean European/African difference favoring Europeans was 80 cm^3. Within race differences, due to method of estimation, averaged 31 cm^3.

Problems of sampling and lack of control over extraneous variables can be cited for many of the individual studies (Tobias, 1970). These difficulties, of course, apply to data from all three racial groups. Body size differences cannot be the cause of the racial differences because Mongoloids have a greater cranial capacity than Negroids although they are often shorter in height and lighter in weight (Eveleth & Tanner, 1990). The racial ordering remains constant even in samples where Negroids are taller than Caucasoids, as in a study of endocranial volume by Simmons (1942) or when the races are statistically equated by adjusting for body size.

As mentioned, Haug (1987) reported a correlation of $r = .479$ ($n = 81$, $p < .001$) between number of cortical neurons and brain volume in cm^3. This translates to Mongoloids, with an average cranial volume of 1,364 cm^3, having 13.767 billion cortical neurons (13.767×10^9). Caucasoids, with a cranial volume of 1,347 cm^3 average 13.665 billion cortical neurons, 102 million less than Mongoloids. Negroids, with a cranial volume of 1,267 cm^3, average 13.185 billion cerebral neurons, 582 million less than Mongoloids and 480 million less than Caucasoids. Most neural tissue goes to maintain bodily functions. Over and above this are "excess neurons" available for general adaptive purposes (Jerison, 1973). However crude the current estimates, hundreds of millions of cerebral cortex neurons differentiate Mongoloids from Negroids (582×10^6 based on those just calculated; 440×10^6 based on those averaged from Tobias as described above). These are probably sufficient to underlie the proportionate achievements in intelligence and social organization.

Brain Size from Infancy to Adolescence

Racial differences in brain size appear early in life. Using data from Broman et al.'s (1987) analyses of the Collaborative Perinatal Project, Rushton (1995, p. 125) calculated that the White children were born with larger heads and larger bodies (a 16 percentile point advantage in each). However, catch-up growth favored Black children in height but not in head perimeter. By age the 4 Black children were 11 percentile points taller than the White children, and by age 7 they were 16 percentile points taller but at age 7 their head perimeter remained 8 percentile points smaller.

Ho, Roessmann, Hause and Monroe (1981) collated brain weights at autopsy from 782 newborns. In absolute terms (unadjusted for other variables), White babies averaged heavier brains than Black babies: 272 grams versus 196 grams. Many of these babies were premature (49% of the White sample and 78% of the Black sample). When the criteria of a gestational age of 38 weeks and a body weight at birth of 2,500 grams was employed to define "full term" in both groups, the racial differences disappeared. Black babies, however, have a biologically based shorter gestation than White babies (Papiernik, Cohen, Richard, de Oca, & Feingold, 1986), and so the appropriateness of imposing these criteria for racial comparisons can be questioned.

Two studies of cranial size have been carried out in adolescents. Lynn (1993) used Lee and Pearson's (1901) equations (1) and (2) with 36 samples of 7- to 15-year-olds gathered by the Philadelphia Growth Center. The core sample consisted of 169 White males, 224 Black males, 135 White females, and 220 Black females. The boys and girls had all been screened for serious illnesses or dental problems and were middle-class from "a solid, stable responsible cross section of the population" (Krogman, 1970, p. 4). After adjusting for the effects of age, stature, and sex, White children averaged 1,250 cm^3 and Black children averaged 1,236 cm^3.

Also with adolescents, Rushton and Osborne (1995) calculated cranial capacities for 187 boys and 285 girls, 222 Whites and 250 Blacks aged 13 to 17 years. Whites averaged 1,269 cm^3 and Blacks 1,251 cm^3. With both Lynn's (1993) and Rushton and Osborne's (1995) adolescent data, however, there is a striking race x sex interaction with the White/Black difference present only for males. Based on the age x sex x race body size data, this appears to be due to maturational differences, with girls maturing earlier than boys and Blacks maturing earlier than Whites, resulting in young Black girls being especially larger in body size relative to their counterparts.

Race Differences in Cognitive Ability

Intelligence Test Scores

Since the time of World War I, when widespread testing began, African descended people have scored lower than Whites on assessments of intelligence and educational attainment (Loehlin, Lindzey, & Spuhler, 1975). Fewer people are aware that Orientals often score higher than Whites on the same tests whether assessed in Canada and the United States, or in their home countries (Vernon, 1982). In an overview of mathematics education, for example, Steen (1987) showed that within the United States, the proportion of Asian American students who achieve high mathematics scores (above 650) on the Scholastic Aptitude Test is twice the national average while the proportion of Black students who do so is less than one fourth the national average.

A review of the international distribution of intelligence test scores has been provided by Lynn (1991b). European populations in North America, Europe, and Australasia average mean IQs of from 85 to 115 with an overall mean of 100 and a standard deviation of 15. That Caucasoids in the United States and Britain obtain virtually identical mean IQs was first demonstrated in a 1932 Scottish survey of 11 year-olds who obtained a mean IQ of 99 on the American Stanford-Binet. Lynn discussed some of the reasons for the variation between European countries, such as sampling accuracy and procedures, and to differences in education and living standards. For example, in the case of children, those in private schools may or may not be included in the samples. The IQs of Indians from the Indian sub-continent and Britain ranged from 85 to 96. A mean of 86 in India was derived from a review by Sinha (1968) of 17 studies of children aged between 9 and 15 years and totalling in excess of 5,000. The ethnic Indians in Britain obtained a mean of 96.

The Mongoloid mean IQs are set out in Table 7. It will be seen that for general intelligence the Mongoloid peoples tend in the majority of studies to obtain somewhat higher means than Caucasoids. This is the case in the United States, Canada, Europe, Japan, Hong Kong, Taiwan, Singapore and The People's Republic of China. The range is from 97 to 110, with a mean of around 106.

A striking feature of the result for Mongoloids is that their verbal IQs are consistently lower than their visuospatial IQs. In most studies the differences are substantial, amounting to between 10 to 15 IQ points. This pattern is present in Japan, Hong Kong, the United States and Canada. This difference also shows up in the United States on the Scholastic Aptitude Test, in which Mongoloids invariably average better than Caucasoids on the mathematics test (largely a measure of general intelligence and visuospatial ability) but also less well than Caucasoids on the verbal test (Wainer, 1988).

Table 7. Mean IQ Scores for Various Mongoloid Samples

Sample	Age	Number	Test	Intelligence			Source
				General	Verbal	Visuo-spatial	
Japan	5-16	1,070	WISC	—	—	103	Lynn, 1977b
Japan	6	240	Vocabulary-spatial	97	89	105	Stevenson et al., 1985
Japan	11	240	Vocabulary-spatial	102	98	107	Stevenson et al., 1985
Japan	4-6	600	WPPSI	103	98	108	Lynn and Hampson, 1988a
Japan	2-8	550	McCarthy	100	92	108	Lynn and Hampson, 1986b
Japan	6-16	1,100	WISC-R	103	101	107	Lynn and Hampson, 1986c
Japan	13-15	178	Differential Aptitude	104	—	114	Lynn, Hampson, and Iwawaki, 1987
Japan	13-14	216	Kyoto NX	101	100	103	Lynn, Hampson, and Bingham, 1987
Japan	3-9	347	Columbia MMS	110	—	—	Misawa et al., 1984
Japan	9	444	Progressive Matrices	110	—	—	Shigehisa and Lynn, 1991
Hong Kong	6-15	4,500	Progressive Matrices	110	—	—	Lynn, Pagliari and Chan, 1988
Hong Kong	10	197	PM, Space Relations	108	92	114	Lynn, Pagliari and Chan, 1988
Hong Kong	9	376	Cattell Culture Fair	113	—	—	Lynn, Hampson, and Lee, 1988
Hong Kong	6	4,858	Coloured P Matrices	116	—	—	Chan and Lynn, 1989
P.R. China	6-14	5,108	Progressive Matrices	101	—	—	Lynn, 1991a
Taiwan	16	1,290	Culture Fair	105	—	—	Rodd, 1959
Singapore	13	147	Progressive Matrices	110	—	—	Lynn, 1977a
Belgium	6-14	19	WISC	110	102	115	Frydman and Lynn, 1989
United States	6-17	4,994	Various	100	97	—	Coleman et al., 1966; Flynn, 1991
United States	6-11	478	Various	101	—	—	Jensen and Inouye, 1980
United States	6-10	2,000	Figure copying	—	—	105	Jensen, 1973
United States	6	80	Hunter Aptitude	106	97	106	Lesser, Fifer and Clark, 1965
United States	6-14	112	Various	107	—	—	Winick et al., 1975
Canada	15	122	Differential Aptitude	105	97	108	Vernon, 1982
Canada	6-8	38	WISC	100	94	107	Kline and Lee, 1972

Source: From R. Lynn (1991b, pp. 264-265, Table 2). Copyright 1991 by The Institute for the Study of Man. Reprinted with permission.

Table 8. Mean IQ Scores for Various Negroid-Caucasoid Mixed Race Samples

Sample	Age	Number	Test	Intelligence			Source
				General	Verbal	Visuo-spatial	
United States	—	—	362 Studies	85	—	—	Shuey, 1966
United States	7	19,000	Wechsler	90	89	93	Broman et al., 1987
United States	2	46	Stanford-Binet	86	—	—	Montie and Fagan, 1988
United States	6-18	4,995	Verbal and non-verbal	84	89	—	Coleman et al., 1966
United States	6	111	WISC	81	86	80	Miele, 1979
United States	6-16	305	WISC-R	84	87	88	Jensen and Reynolds, 1982
United States	7-14	642	PMA	77	77	83	Baughman and Dahlstrom, 1968
United States	6-11	2,518	Various	84	—	—	Jensen and Inouye, 1980
S. Africa coloured	10-14	4,721	Army Beta	84	—	—	Fick, 1929
Barbados	9-15	108	WISC-R	82	84	84	Galler et al., 1986
Britain	11	113	NFER	86	87	—	Mackintosh and Mascie-Taylor, 1985
Britain	10	125	British Ability Scales	94	92	—	Mackintosh and Mascie-Taylor, 1985
Britain	8-12	205	NFER	87	—	—	Scarr et al., 1983
Jamaica	10-11	50	Various	75	82	90	Vernon, 1969
Jamaica	11	1,730	Moray House	72	72	—	Manley, 1963; Vernon, 1969
Jamaica	5-12	71	WISC	66	74	64	Hertzig et al., 1972

Source: From R. Lynn (1991b, p. 269, Table 4). Copyright 1991 by The Institute for the Study of Man. Reprinted with permission.

Research on the academic accomplishments of Mongoloids in the United States continues to grow. Caplan, Choy and Whitmore (1992) gathered survey and test score data on 6,750 Indochinese refugees in five urban areas around the United States. Unlike some of the previously studied populations of "boat people," these refugees had had limited exposure to Western culture, knew virtually no English when they arrived and often had a history of physical and emotional trauma. Often they came with nothing more than the clothes they wore. All the children attended schools in low-income metropolitan areas. The results showed that whether measured by school grades or nationally normed standardized tests, the children were above average overall, "spectacularly" so in mathematics.

The mean IQs of Negroids are invariably found to be lower than those of Caucasoids. Three hundred and sixty-two investigations done in the United States were presented by Shuey (1966) who reported the overall mean IQ of African Americans to be approximately 85. Subsequent studies in the United States such as those by Coleman et al. (1966), Broman et al. (1987) and others have confirmed this figure. Many of these studies are shown in Table 8. For the United States, seven major post Shuey (1966) studies were chosen because of their special interest by virtue of the large number of subjects, because they yield IQs for the verbal and visuospatial abilities, or because they are derived from young children. These show that the Negroid mean IQ of approximately 85 is present among children as young as two to six year olds. In Britain, three studies of Afro-Caribbeans obtained mean IQs of 86, 94, and 87, broadly similar to those in the United States. Figures are available for two of the Caribbean islands, namely Barbados (mean IQ = 82) and Jamaica (mean IQ = 66-75).

As a result of these studies, carried out across different intelligence tests and cohorts, it is sometimes assumed that the mean IQ of all Negroids is approximately 85. Lynn noted, however, that most African Americans are Negroid-Caucasoid hybrids with about 25 percent Caucasian admixture (Chakraborty, Kamboh, Nwankwo, & Ferrell, 1992) and he believed a similar proportion was probably true of Blacks in the West Indies and Britain. It is possible, therefore, that the mean IQs of non-mixed Africans will be lower than that of the hybrids. Lynn tested this hypothesis by examining the literature from Africa (see Table 9).

An early study of the intelligence of "pure" African Negroids was carried out in South Africa by Fick (1929). He administered the American Army Beta Test, a non-verbal test designed for those who could not speak English, to 10-14 year old White, Black African and Coloured (mainly Negroid-Caucasoid hybrid) school children. In relation to the White mean of 100, based on more than 10,000 children, largely urban Black African children obtained a mean IQ of 65, while urban Coloured children obtained a mean IQ of 84. Thus South African Coloureds obtained a mean IQ virtually identical to that of African Americans.

Table 9. Mean IQ Scores for Various African Negroid Samples

Sample	Age	Number	Test	Intelligence			Source
				General	Verbal	Visuo-spatial	
Congo	adults	320	Progressive Matrices	65	—	—	Ombredane et al., 1952
Ghana	adults	225	Culture Fair	80	—	—	Buj, 1981
Nigeria	6-13	87	Colored Matrices, PMA	75	—	81	Fahrmeier, 1975
Nigeria	adults	—	Progressive Matrices	86	—	—	Wober, 1969
South Africa	8-16	1,220	Progressive Matrices	81	—	—	Notcuff, 1950
South Africa	adults	703	Progressive Matrices	75	—	—	Notcuff, 1950
South Africa	10-14	293	Army Beta	65	—	—	Fick, 1929
South Africa	9	350	Progressive Matrices	67	—	—	Lynn and Holmshaw, 1990
South Africa	16	1,093	Junior Aptitude	69	60	69	Owen, 1989
Uganda	12	50	Various	80	—	—	Vernon, 1969
Zambia	adults	1,011	Progressive Matrices	75	—	—	Pons, 1974; Crawford Nutt, 1976

Source: Lynn (1991b, p. 267, Table 3).

The other studies of the IQs of Black Africans summarized in Table 9 show means in the range of 65 to 81. Lynn cited the work of Owen (1989) as the best single study. Owen presented results for 1,093 16 year olds in the eighth grade who had been in school for around eight years and should have been knowledgable about paper and pencil tests. The test used was the South African Junior Aptitude which provides measures of verbal and non verbal reasoning, spatial ability, verbal comprehension, perceptual speed and memory. The mean IQ of the sample in comparison with White South African norms is 69 which is also around the median of the studies listed in Table 9. Lynn rounded this figure to 70 and took it as the approximate mean for pure Negroids.

Owen (1992) published the South African study cited in Table 9 more fully. The scores of high school students on Raven's Standard Progressive Matrices showed clear racial differences with 1,065 Whites = 45.27 (SD = 6.34), 1,063 East Indians = 41.99 (SD = 8.24), 778 Coloureds = 36.69 (SD = 8.89) and 1,093 pure Negroids = 27.65 (SD = 10.72). Thus Negroids are from 1.5 to 2.7 standard deviations below the two Caucasoid populations and about 1 standard deviation lower than the Coloureds. The four groups showed little difference in test reliabilities, the rank order of item difficulties, item discrimination values, and the loadings of items on the first principal component. Owen (1992, p. 149) concluded: "Consequently, from a psychometric point of view, the [test] is not culturally biased."

Two other published studies have supported Lynn's (1991) estimate of a Black African IQ of 70. Zindi (1994) administered the Wechsler Test to a representative sample of children in Zimbabwe and Lynn (1994a) examined Ethiopian immigrants into Israel. Language was probably not the mediator in the Zimbabwe study because the racial disparity was as large on "Performance" and "Culture-Fair" tests as on "Verbal" tests.

Parenthetically, we note that Lynn (1991b) also summarized the results of studies on the intelligence of Amerindians. The largest study is that of Coleman et al (1966) which obtained a mean of 94, but a number of studies have reported means in the 70 to 90 range. The median of 15 studies Lynn reviewed was 89 which Lynn took as a reasonable approximation, indicating that the Amerindian mean IQ falls someway between that of Caucasoids and Negroid-Caucasoid hybrids. The same intermediate position is occupied by Amerindians in performance on the Scholastic Aptitude Test (Wainer, 1988). In addition, all the studies of Amerindians have found that they have higher visuospatial than verbal IQs. The studies listed are those where the Amerindians speak English as their first language, so this pattern of results is unlikely to be solely due to the difficulty of taking the verbal tests in an unfamiliar language. The verbal-visuospatial disparity is also picked up in the Scholastic Aptitude Test, where Amerindians invariably score higher on the mathematical test than on the verbal (Wainer, 1988).

Finally, Lynn examined the published IQ scores for several South East Asian peoples, including Polynesians, Micronesians, Melanesians, Maoris and Australian Aborigines. Apart from the low mean of 67 for a small sample of Australian Aborigine children, all the mean IQs lie in the range of 80-95. The one study to include measures of general, verbal and visuospatial abilities for New Zealand Maoris shows that this population does not share the strong visuospatial-weak verbal ability profile of Mongoloids and Amerindians. Although the intelligence of this group of peoples has not been extensively researched, Lynn suggested there are sufficient studies to suggest a mean IQ of about 90.

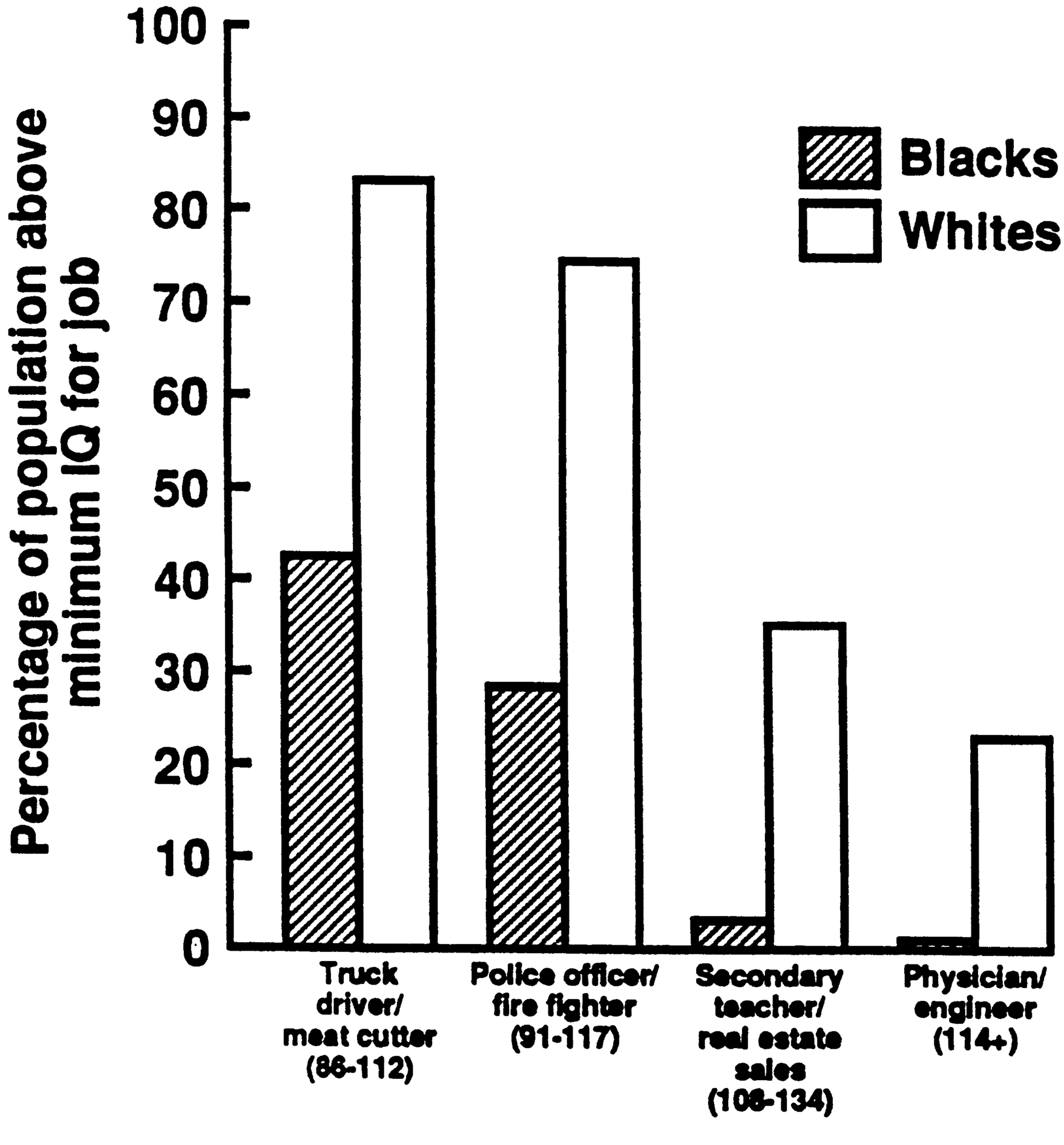
Decision Times

Speed of decision processing on elementary cognitive tasks typically shows the same three-way racial pattern as do test scores. Investigations have been done on 9- to 12-year-olds from six countries. Children were asked to decide which of several lights stands out from others and move a hand to press a button. All children can perform the task in less than 1-second, but children with higher IQ scores perform faster (after controlling for movement time) than do those with lower scores. Lynn (1991b) found that Asian children from Hong Kong and Japan were faster than were European children from Britain and Ireland, who in turn were faster than African children from South Africa (see also Lynn & Shigehisa, 1991). Using similar tasks, as well as those involving retrieval of well-learned facts from long-term memory, this pattern of racial differences was also found in California (Jensen, 1993; Jensen & Whang, 1993, 1994).

Occupational Achievement

Gottfredson (1986, 1987) has suggested that occupations be viewed as analogous to differentially *g*-loaded mental tests. Large-scale studies from World War I through the 1980s have shown that occupations differ considerably in the mean intelligence levels of their incumbents. The mean level of intelligence of the occupation, in turn, correlates highly with the prestige level of the occupation. Gottfredson reasoned that the overall intellectual complexity of the work should effect the percentage of workers who are Black. Figure 2 provides data relevant to this conjecture.

Gottfredson first determined the IQ ranges from which workers have most often been recruited to different occupations. Then she used nationally representative mental test data to determine the proportions of Blacks and Whites falling within each of those IQ recruitment ranges. Third she computed the ratio of Blacks to Whites who would be eligible for each occupation on the basis of intelligence alone. Black-White parity in employment is represented by a ratio of 1.00. She found that the ratios ranged from 0.72 for truck drivers



Note: The lower average distribution of IQ among Blacks leads to a disproportionate underrepresentation in occupations selected on the basis of hight IQ.

Source: Drawn from data in Gottfredson (1986, 1987). From *Race, Evolution, and Behavior* (p. 145, Figure 6.3) by J.P. Rushton, 1995, New Brunswick, NJ: Transaction. Copyright 1995 by Transaction Publishers. Reprinted with permission.

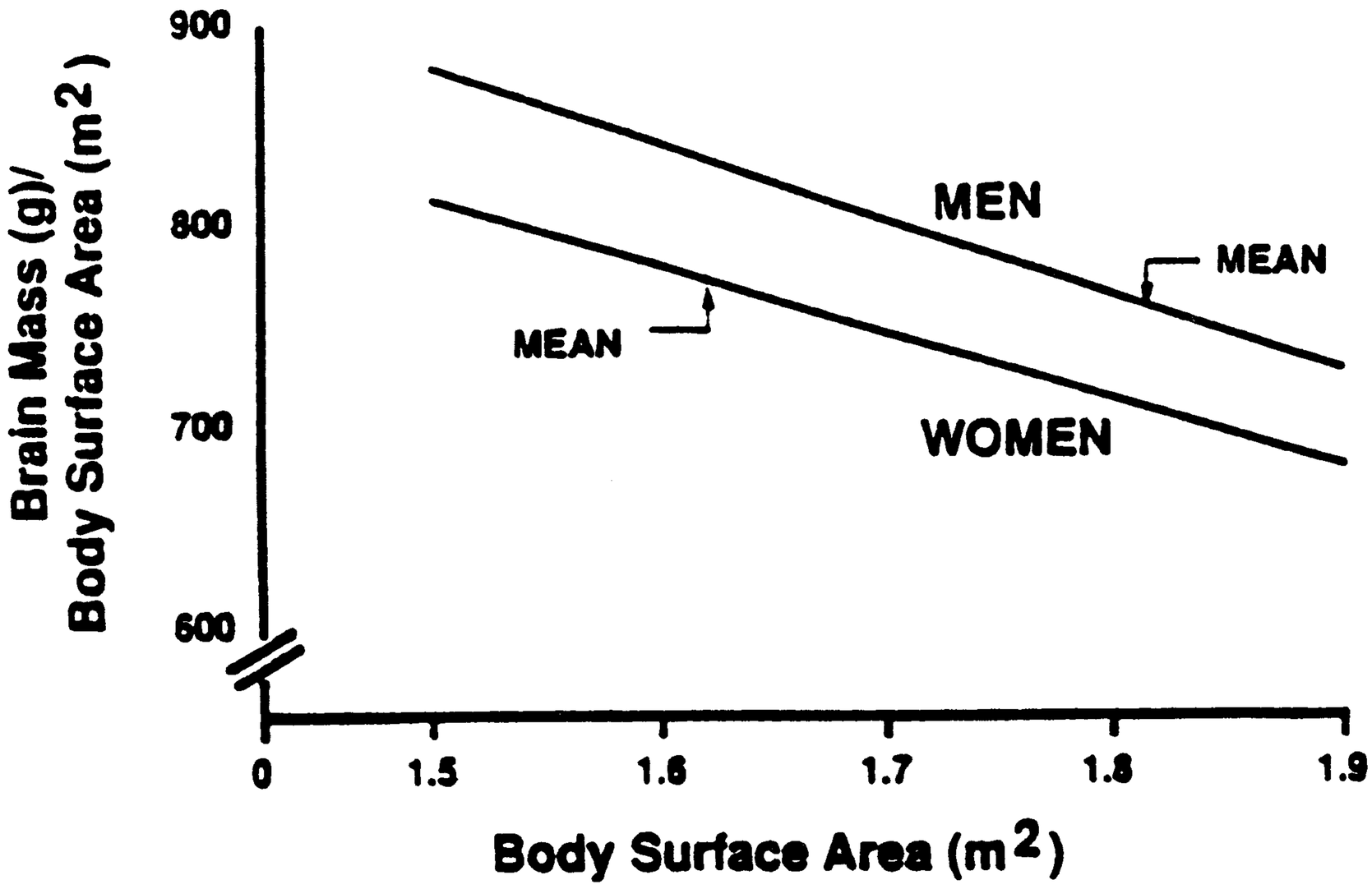
Figure 2. Percentage of Blacks and Whites in the United States above Minimum IQ Required for Various Occupations

to 0.05 for physicians, proportional to the actually observed Black/White ratios of 0.98 to 0.30. Note that observed Black-White differences in employment are smaller than those expected on the basis of intelligence alone, a finding consistent with data showing that mean IQs are lower for Blacks than for Whites in the same occupational category and for Black versus White applicants for the same jobs.

Sex Differences in Brain Size

The absolute difference in brain size between men and women, so clearly observable throughout Table 6, has not been disputed since at least the time of Broca (1861). It is often claimed, however, that the sex difference disappears when corrections are made for body size or age of people sampled (Gould, 1981; Lewontin et al., 1984). Regardless, a recent study by Ankney (1992) demonstrated that the sex difference in brain size remains after correction for body size in a sample of similarly aged men and women (following tentative results by Dekaban & Sadowsky, 1978; Gur et al., 1991; Holloway, 1980; Swaab & Hofman, 1984; Hofman & Swaab, 1991; Willerman et al. 1991).

Ankney (1992) argued that the large sex difference in brain size went unnoticed for so long because earlier studies used improper statistical techniques to correct for sex differences in body size and, thus, incorrectly made a large difference "disappear." The serious methodological error was the use of brain-mass/body-size ratios instead of analysis of covariance (see Packard & Boardman, 1988). Ankney (1992) illustrated why this is erroneous



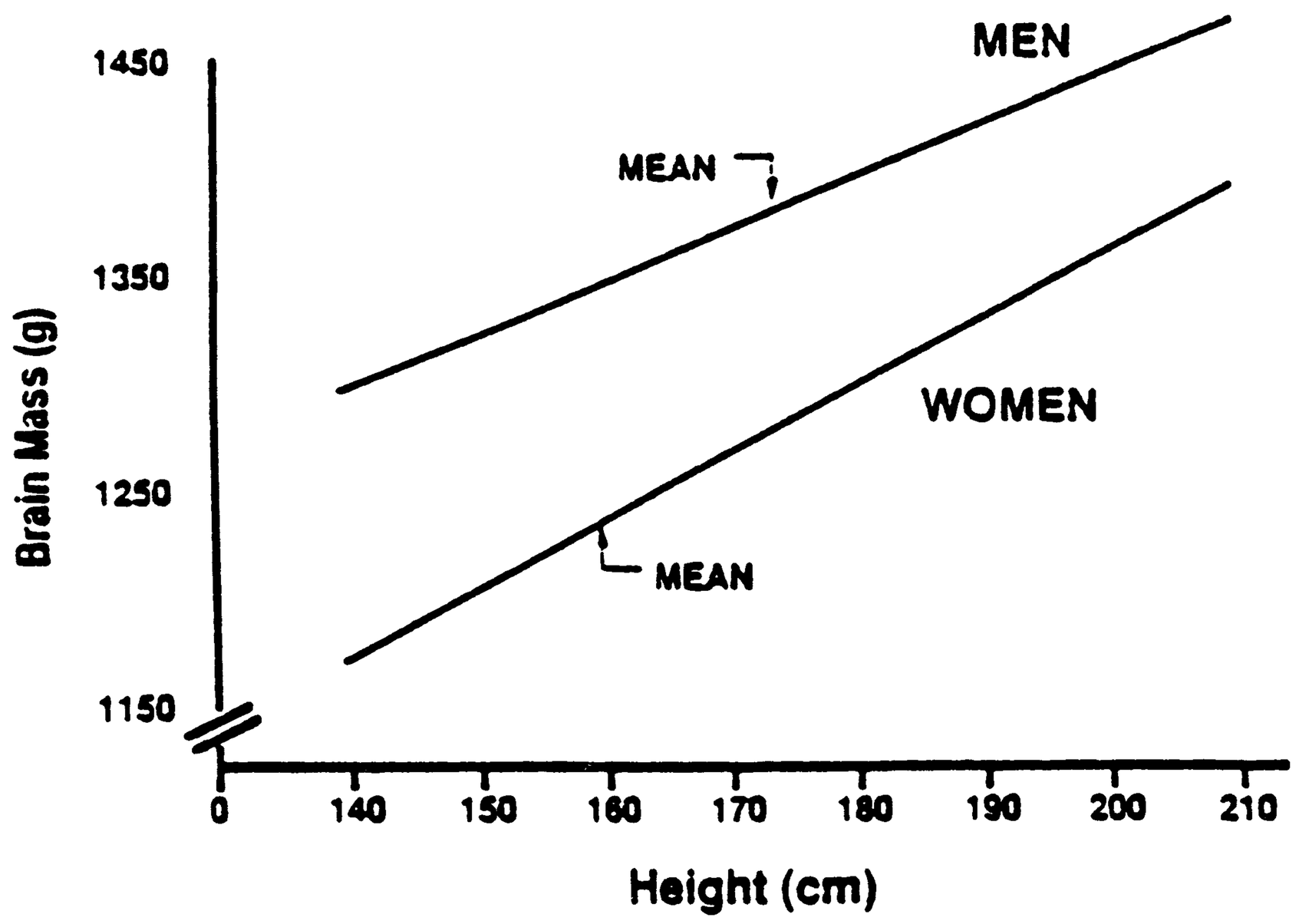
Notes: The ratios were calculated by estimating brain mass at a given body surface area using the equations in Ho et al. (1980b, Table 3): *Men*, brain mass = 1077 *g* (±56) + 173 (±31) x body surface areas ($r = .27$, $p < .01$); *Women*, brain mass = 949 *g* (±52) + 188 (±32) x body surface area ($r = .24$, $p < .01$).

Source: From Ankney (1992, p. 331, Figure 1). Copyright 1992 by Ablex Publishing Corporation. Reprinted with permission.

Figure 3. The Relation Between the Ratio of Brain-Mass/Body-Surface-Area and Body-Surface-Area in White Men and Women

by showing that, in both men and women, the ratio of brain-mass to body-size declines as body-size increases. Thus, as can be seen in Figure 3, larger women have a lower ratio than do smaller women as do larger men compared to smaller men. Therefore, because the average-sized man is larger than the average-sized woman, their brain-mass to body-size ratios are similar (Figure 3). Consequently, the only meaningful approach is to compare brain-mass to body-size ratios of men and women of equal size. Such comparisons show that at any given size, the ratio of brain-mass to body-size is much higher in men than in women (see Figure 3).

Ankney reexamined autopsy data on 1,261 American adults (Ho et al., 1980a, 1980b) and found that at any given body-surface-area or height, brains of White men are heavier than those of White women as are brains of Black men heavier than those of Black women. For example, among 168 cm tall Whites (the approximately overall mean height for men and women combined), brain mass of men averages about 100 grams heavier than that of women (Figure 4), whereas the average difference in brain mass, uncorrected for body size, was 140 g. Thus, only about 30 percent of the sex difference in brain-size is due to differences in body-size.



Notes: Lines drawn from equations in Ho et al. (1980b, Table 1): *Men*, brain mass = 920 g (± 113) + 2.70 (± 65) x body height ($r = .20$, $p < .01$); *Women*, brain mass = 748 g (± 104) + 3.10 (± .64) x body height ($r = .24$, $p < .01$).

Source: From Ankney (1992, p. 333, Figure 4). Copyright 1992 by Ablex Publishing Corporation. Reprinted with permission.

Figure 4. The Relation Between Brain Mass and Body Height in White Men and Women

Ankney's results were confirmed in the study of cranial capacity from a stratified random sample of 6,325 U.S. Army personnel shown earlier in Table 4. After adjusting, via analysis of covariance, for effects of age, stature, weight, military rank and race, men averaged 1,442 cm^3 and women 1,332 cm^3. This difference was found in all of the 20 or more separate analyses done to rule out any body size effect (Rushton, 1992a, pp. 406-408). Moreover, the difference was replicated across samples of Asians, Whites, and Blacks, as well as across officers and enlisted personnel. The sex difference of 110 cm^3 found by Rushton, from analysis of external head measurements, is remarkably similar to that (100 g) obtained by Ankney, from analysis of brain mass (1 $cm^3 = 1.036$ g; Hofman, 1991). Other studies have confirmed the sex difference. Andreasen (1993) corroborated the sex difference in adult brain size using MRI (see also Gur et al. 1991; Harvey et al., 1994; Willerman et al., 1991).

From birth through early months, we found the sex difference held in several autopsy studies when, following Ankney's (1992) procedure (see Figure 4), we compared brain masses of boys and girls after matching them for stature (Dekaban & Sadowsky, 1978; Pakkenberg & Voigt, 1964; Voigt & Pakkenberg, 1983). In children from 4 to 7 years of age, sex differences are found with brain size inferred from external head measurements. After adjustments for body size and race, sex differences in head perimeter are about 0.40 S.D. (Jensen & Johnson, 1994). From 7 to 17 years, sex differences in cranial capacity are in the range of 60 to 100 cm^3 (Lynn, 1993; Rushton & Osborne, 1995).

Sex Differences in Cognitive Ability

These results present a paradox. Women have proportionately smaller brains than do men, but apparently have the same intelligence test scores. According to Kimura (1992), however, women excel in verbal ability, perceptual speed, and motor coordination within personal space, whereas men do better on various spatial tests and on tests of mathematical reasoning. Although controversy exists about the magnitude of the sex difference in spatial ability under various testing conditions, reviews by Pool (1994) and Voyer, Voyer and Bryden (1995) show that on the "purest" spatial measures, such as rotating an imaginary object, or shooting at a moving rather than a stationary target, the sex difference approaches 1 standard deviation. Thus Ankney (1992) hypothesized that the sex difference in brain size relates to those intellectual abilities at which men excel, that is, those abilities that require more "brain power." Analogously, increasing word processing power in a computer requires some extra capacity, but increasing 3-dimensional processing, as in graphics, requires a major increase in capacity.

Consistent with Ankney's hypothesis, Andreasen et al. (1993) used MRI to estimate brain volume in 37 men and 30 women and showed that brain size correlated more highly with Performance IQ than with Verbal IQ in men

($r = 0.43$ vs 0.33), and that the opposite pattern occurred in women ($r = 0.30$ vs 0.43), that is, increased brain size tended to be more strongly related to Performance IQ than to Verbal IQ in men and vice versa in women. Similarly, Wickett et al. (1994) used MRI to measure brain volume in 40 women and found that it correlated 0.44 with Verbal IQ, but only 0.28 with Performance IQ. Moreover, although Lynn (1994b) found that men averaged 4 points higher than did women on standard IQ tests, Ankney (1995) showed that nearly all of this difference derived form men's higher scores on spatial and mathematical reasoning subtests. As noted by Ankney (1995), what is truly needed to test his hypothesis are the slopes, not just the correlations, from the regressions of various mental abilities and IQ on brain size for men and women.

The nineteenth century proposition that men average slightly higher in general intelligence than do women (e.g., Broca, 1861, p. 153) has also been reactivated. Lynn's (1994b) resolution of the sex difference in brain size paradox was to contradict (with evidence) the consensus that there is no difference in general intelligence. He reviewed data from Britain, Greece, China, Israel, the Netherlands, Norway, Sweden and Indonesia, as well as the United States, to show that men averaged about 4 IQ points more than did women on a number of published intelligence tests. Lynn also found men averaged better on high performance school examinations such as the Advanced level examination used for entrance to university in Britain as well as in final grades at British universities. In the United States, Lynn showed men get higher scores on both the Verbal and the Quantitative components of the Scholastic Aptitude Test (SAT).

Independently, Jackson (1993, 1995) found the sex differences were mainly a matter of *g*. On the Scottish standardization of the Wechsler Intelligence Scale for Children-Revised, based on the responses of 1,400 children aged 6 to 16 years, boys scored higher on highly *g*-loaded tests and girls on less highly *g*-loaded tests. With 94,970 male and 88,632 female medical school applicants in Germany over an 11 year period, males consistently scored higher on highly *g*-loaded tests and females higher on less highly *g*-loaded tests, the difference being 0.30 of a standard deviation or about 12 percentile units at the mean. From the same data set, Stumpf and Jackson (1994) reported a half standard deviation advantage to men in general reasoning ability. Finally, with 112,516 male and female respondents on the SAT, factor analysis of the items showed an 8-10 percentile point advantage to men on the *g* factor for all comparisons undertaken, that is, for groups of males and females differentiated for family income, mother's education, father's education, and ethnic background, regardless of whether the *g* factor was estimated from verbal, quantitative, or total sets of items.

Origins

Behavioral Genetics

A first study of the genetic contribution to cranial size, and by inference to brain size, has been made in a study of 236 pairs of adolescent twins (472 individuals) aged 13 to 17 years, White and Black, male and female (Rushton & Osborne, 1995). Heritability for the total sample ranged from 38 percent to 51 percent, depending on particular adjustments made for body size. Environmental effects common to both twins, like parental socioeconomic status, ranged from 6 percent to 20 percent and environmental effects unique to each twin, like illness or trauma, ranged from 42 percent to 52 percent. The heritability estimates did not vary systematically by sex or race, although there was a trend for them to be lower in Blacks than in Whites.

Heritabilities for mental ability range from 50 percent to 80 percent and have been established in numerous adoption, twin, and family studies (Bouchard & McGue, 1981). Noteworthy are the 80 percent heritabilities found in adult twins reared apart (Bouchard, Lykken, McGue, Segal, & Tellegen, 1990; Pedersen, Plomin, Nesselroade, & McClearn, 1992). Genetic influence is also found in studies of non-Whites, including African Americans (Osborne, 1980; Scarr, Weinberg, & Waldman, 1993) and Japanese (Lynn & Hattori, 1990).

Findings such as these led Sandra Scarr to title her 1986 presidential address to the Behavior Genetics Association, "Three Cheers for Behavioral Genetics." She observed that "the war [between nature and nurture] is largely over." Scarr accepted that genetics underlay existing White social class differences in mental ability in the United States and Europe, although this may not have been the case for earlier generations when social mobility was more restricted. Because racial barriers are less permeable than class barriers, Scarr (1987) interpreted her own research of 7-year-old Black and mixed-race children adopted into White middle-class families as showing an environmental cause of racial differences.

Other transracial adoption studies, including a follow-up to Scarr's 7-year-old Black children to when they were 17, indicate a genetic contribution to cognitive ability. Studies of Korean and Vietnamese children adopted into White American and White Belgian homes showed that, although as babies many had been hospitalized for malnutrition, they grew to excel in academic ability with IQs 10 points or more higher than their adoptive national norms (Clark & Hanisee, 1982; Frydman & Lynn, 1989; Winick, Meyer, & Harris, 1975). By contrast, Weinberg, Scarr, and Waldman (1992) found that at age 17, Black and mixed-race children adopted into White middle-class families performed at a lower level than did White siblings with whom they had been raised. Adopted White children had an average IQ of 106, an average aptitude based on national norms at the 59th percentile, and a class rank at the 54th percentile; Mixed-race children had an average IQ of 99, an aptitude at the

53rd percentile, and a class rank at the 40th percentile; and Black children had an average IQ of 89, an aptitude at the 42nd percentile, and a class rank at the 36th percentile.

Multifarious other sources of evidence suggest that racial differences in intelligence are substantially genetic. For example, Black-White differences are most pronounced on more *g*-loaded subtests; that is, on the general factor common to diverse cognitive tests (Jensen, 1985). The *g*-loadings are correlated with a number of biological variables including brain evoked potentials, heritability coefficients determined from twin studies, and the degree to which children's test scores are depressed by inbreeding and raised by outbreeding (Jensen, 1997). Also, genetic weights established from inbreeding depression studies in the Japanese population directly predict the magnitude of the Black-White differences on the various subtests of the Wechsler Intelligence Scale for Children (Rushton, 1989b). There is no other explanation for inbreeding depression than a genetic one.

Evolution

Because of a three-fold increase in relative size of the hominid brain over the last 3 million years in which australopithecenes averaged 500 cm^3 (the size of a chimpanzee brain), *Homo erectus* about 1,000 cm^3, and *Homo sapiens*, about 1,300 cm^3, it is reasonable to hypothesize that bigger brains evolved via natural selection for increased intelligence (Jerison, 1973). Metabolically, the human brain is an expensive organ. Representing only 2 percent of body mass, the brain uses about 5 percent of basal metbabolic rate in rats, cats, and dogs, about 10 percent in rhesus monkeys and other primates, and about 20 percent in humans (Armstrong, 1990). Across species, large brains are related to other life history traits, such as a longer gestation, a slower rate of maturation, a higher offspring survival, a lower reproductive output, and a longer life (Pagel & Harvey, 1988; Hofman, 1993). From an adaptationist perspective, unless large brains substantially contributed to evolutionary fitness (defined as increased survival of genes through successive generations), they would not have evolved.

Evolutionary hypotheses for why Asians average the largest brains and have the most intelligence have been provided (Rushton, 1995). The currently accepted view of human origins, the "African Eve" theory, posits a beginning in Africa some 200,000 years ago, an exodus through the Middle East with an African/non-African split about 110,000 years ago, and a Caucasoid/Mongoloid split about 40,000 years ago. Evolutionary selection pressures were different in the hot savanna where Africans evolved than in the cold arctic where Mongoloids evolved. The further north out of Africa that populations migrated, the more they encountered the cognitively demanding problems of gathering and storing food, gaining shelter, making clothes, and raising

children during prolonged winters. Rushton (1995) proposed that as the original African populations evolved into Caucasoids and Mongoloids, they did so in the direction of larger brains, slower rates of maturation, and other traits differentiating the populations.

The sexual dimorphism in cranial size and cognitive ability likely originated partly through evolutionary selection of men's hunting ability (Kolakowski & Malina, 1974; Ankney, 1992) and partly through the reproductive success socially dominant men have traditionally enjoyed (Lynn, 1994b). The brain size of individuals, of course, is also affected by nutrition and experience, most obviously through illness and trauma.

CONCLUSION

Differences in brain size are correlated with differences in cognitive abilities, and both brain size and cognitive ability are correlated with sex and race. Correlation does not prove cause and effect, but, just as zero correlations provide no support for a hypothesis of cause and effect, non-zero correlations do provide support. We are convinced that the brain-size/cognitive-ability correlations that we have reported are, in fact, due to cause and effect. This is because we are unaware of any variable, other than the brain, that can directly mediate cognitive ability. Of course, brain size is not the only mediator of brain function; Miller (1994) reviewed data suggesting that amount of brain myelination is related to IQ. Regardless, we believe that the important research questions are (1) what is responsible for these group differences in brain size, that is, are they genetically and/or environmentally caused?, and (2) why does variation in brain size correlate with differences in cognitive abililty?

Numerous problem areas remain to be researched. For example, it is not known whether women have fewer neurons than do men; there may be greater cortical packing density in women, and thus, it is myelin thickness or some other variable that is responsible for the sex differences in brain size (Haug, 1987). In a post-mortem study of brain tissue from the temporal lobes of five women and four men, Witelson, Glezer and Kigan (1995) supported the hypothesis that women's neurons are packed more tightly. It is unknown, however, whether tightly packed neurons are more or less efficient than are those more widely spaced; the latter may allow a greater number of synaptic connections. Serious paradoxes also require resolution. For example, White women have brain sizes equal to or smaller than do Black men, but nonetheless, score higher than do Black men in general cognitive ability. Additional research using magnetic resonance imaging or behavior genetic techiques is certain to enrich knowledge of these important relationships. MRI may identify features of the brain that correlate even more highly with IQ than does volume. Some possibilities are neuronal density, white/gray contrast, ventricle/brain ratio,

and various specific brain regions. More generally, as Broca and other nineteenth century scientists conjectured so long ago, it may be the complexities of the convolutions of the brain, and the varieties and efficiencies of its commissures, rather than its actual size, that is related to intellectual ability and that differentiates populations.

REFERENCES

Andreasen, N.C. (1993, November). Sex differences in the brain: Perspectives from neuroimaging. Presented at the Society for Neuroscience, Washington, DC.

Andreasen, N.C., Ehrhardt, J.C., Swayze, V.W., Alliger, R.J., Yuh, W.T.C., Cohen, G., & Ziebell, S. (1990). Magnetic resonance imaging of the brain in schizophrenia. *Archives of General Psychiatry, 47*, 35-44.

Andreasen, N.C., Flaum, M., Swayze, V., O'Leary, D.S., Alliger, R., Cohen, G., Ehrardt, J., & Yuh, W.T.C. (1993). Intelligence and brain structure in normal individuals. *American Journal of Psychiatry, 150*, 130-134.

Ankney, C.D. (1992). Sex differences in relative brain size: The mismeasure of woman, too? *Intelligence, 16*, 329-336.

Ankney, C.D. (1995). Sex differences in brain size and mental abilities: Comments on R. Lynn and D. Kimura. *Personality and Individual Differences, 18,* 423-424.

Appel, F.W., & Appel, E.M. (1942). Intracranial variation in the weight of the human brain. *Human Biology, 14*, 235-250.

Armstrong, E. (1990). Brains, bodies and metabolism. *Brain, Behavior and Evolution, 36*, 166-176.

Baker, J.R. (1974). *Race*. Oxford: Oxford University Press.

Baughman, E.E., & Dahlstrom, W.G. (1968). *Negro and White Children.* New York: Academic Press.

Beals, K.L., Smith, C.L. & Dodd, S.M. (1984). Brain size, cranial morphology, climate, and time machines. *Current Anthropology, 25*, 301-330.

Bean, R.B. (1906). Some racial peculiarities of the Negro brain. *American Journal of Anatomy, 5*, 353-432.

Bogaert, A.F., & Rushton, J.P. (1989). Sexuality, delinquency and r/K reproductive strategies: Data from a Canadian university sample. *Personality and Individual Differences, 10*, 1071-1077.

Bouchard, T.J. Jr., & McGue, M. (1981). Familial studies of intelligence: A review. *Science, 212*, 1055-1059.

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

Broca, P. (1861). Sur le volume et la forme du cerveau suivant les individus et suivant les races. *Bulletin SociÄtÄ d'Anthropologie Paris, 2*, 139-207, 301-321, 441-446.

Broman, S.H., Nichols, P.L., Shaughnessy, P., & Kennedy, W. (1987). *Retardation in young children.* Hillsdale, NJ: Erlbaum

Buj, V. (1981). Average IQ values in various European countries. *Personality and Individual Differences, 2*, 168-169.

Caplan, N., Choy, M.H., & Whitmore, J.K. (1992). Indochinese refugee families and academic achievement. *Scientific American, 266*, 18-24.

Cattell, R.B. (1982). *The inheritance of personality and ability.* New York: Academic Press.

Chakraborty, R., Kamboh, M.I., Nwankwo, M., & Ferrell, R.E. (1992). Caucasian genes in American blacks. New data. *American Journal of Human Genetics, 50*, 145-155.

Chan, J., & Lynn, R. (1989). The intelligence of 6-year-olds in Hong Kong. *Journal of Biosocial Science, 21*, 461-464.

Clark, E.A., & Hanisee, J. (1982). Intellectual and adaptive performance of Asian children in adoptive American settings. *Developmental Psychology, 18*, 595-599.

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPortland, J., Mood, A.M., Weinfeld, F.D., & York, R.L. (1966). *Equality of educational opportunity*, 2 vols. Washington, DC: U.S. Office of Education.

Coon, C.S. (1982). *Racial adaptations*. Chicago: Nelson-Hall.

Crawford Nutt (1976). *African IQ in Zambia*. Cited in R.Lynn (1991b).

Dekaban, A.S., & Sadowsky, D. (1978). Changes in brain weights during the span of human life: Relation of brain weights to body heights and body weights. *Annals of Neurology, 4*, 345-356.

DeMyer, M.K., Gilmor, R.L., Hendrie, H.C., DeMyer, W.E., Augustyn, G.T., & Jackson, R.K. (1988). Magnetic resonance brain images in schizophrenic and normal subjects: Influence of diagnosis and education. *Schizophrenia Bulletin, 14*, 21-32.

Egan, V., Wickett, J.C., & Vernon, P.A. (1995). Brain size and intelligence: Erratum, addendum, and correction. *Personality and Individual Differences, 19*, 113-115.

Estabrooks, G.H. (1928). The relation between cranial capacity, relative cranial capacity and intelligence in school children. *Journal of Applied Psychology, 12*, 524-529.

Eveleth, P.B., & Tanner, J.M. (1990). *Worldwide variation in human growth* (2nd ed.). London: Cambridge University Press.

Fahrmeier, E.D. (1975). The effect of school attendance on intellectual development in northern Nigeria. *Child Development, 46*, 281-285.

Fick, M.L. (1929). Intelligence test results of poor white, native (Zulu), coloured and Indian school children and the educational and social implications. *South African Journal of Science, 26*, 904-920.

Fisch, R.O., Bilek, M.K., Horrobin, J.M., & Chang, P.N. (1976). Children with superior intelligence at 7 years of age. *American Journal of Diseases in Children, 130*, 481-487.

Flaum, M., Andreasen, N.C., Swayze II, V.W., O'Leary, D.S., & Alliger, R.J. (1994). IQ and brain size in schizophrenia. *Psychiatry Research, 53*, 243-257.

Flynn, J.R. (1991). *Asian Americans: Achievement beyond IQ*. Hillsdale, NJ: Erlbaum.

Frydman, M., & Lynn, R. (1989). The intelligence of Korean children adopted in Belgium. *Personality and Individual Differences, 10*, 1323-1326.

Galler, J.R., Ramsey, F., & Forde, V. (1986). A follow up study in the influence of early malnutrition on subsequent development. *Nutrition and Behavior, 3*, 211-222.

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

Gottfredson, L.S. (1986). Societal consequences of the g factor in employment. *Journal of Vocational Behavior, 29*, 379-410.

Gottfredson, L.S. (1987). The practical significance of black-white differences in intelligence. *Behavioral and Brain Sciences, 10*, 510-512.

Gould, S.J. (1978). Morton's ranking of races by cranial capacity. *Science, 200*, 503-509.

Gould, S.J. (1981). *The mismeasure of man*. New York: Norton.

Gur, R.C., Mozley, P.D., Resnick, S.M., Gottlieb, G.L., Kohn, M., Zimmerman, R., Herman, G., Atlas, S., Grossman, R., Berretta, D., Erwin, R., & Gur, R.E. (1991). Gender differences in age effect on brain atrophy measured by magnetic resonance imaging. *Proceedings of the National Academy of Sciences, 88*, 2845-2849

Hack, M., Breslau, N., Weissman, B., Aram, D., Klein, N., & Borawski, E. (1991). Effect of very low birth weight and subnormal head size on cognitive abilities at school age. *New England Journal of Medicine, 325*, 231-237.

Haier, R.J., Chueh, D., Touchette, P., Lott, I., Buchsbaum, M.S. Macmillan, D., Sandman, C., LaCasse, L., & Sosa, E. (1995). Brain size and cerebral glucose metabolic rate and nonspecific mental retardation and Down syndrome. *Intelligence, 20*, 191-210.

Harvey, I., Persaud, R., Ron, M.A., Baker, G., & Murray, R.M. (1994). Volumetric MRI measurements in bipolars compared with schizophrenics and healthy controls. *Psychological Medicine, 24*, 689-699.

Haug, H. (1987). Brain sizes, surfaces, and neuronal sizes of the cortex cerebri: A stereological investigation of man and his variability and a comparison with some species of mammals (primates, whales, marsupials, insectivores, and one elephant). *American Journal of Anatomy, 180*, 126-142.

Henneberg, M., Budnik, A., Pezacka, M., & Puch, A.E. (1985). Head size, body size and intelligence: intraspecific correlations in Homo sapiens sapiens. *Homo, 36*, 207-218.

Herskovits, M.J. (1930). *The anthropometry of the American Negro*. New York: Columbia University Press.

Hertzig, M.E., Birch, H.G., Richardson, S.A., & Tizard, J. (1972). Intellectual levels of school children severely malnourished during the first two years of life. *Pediatrics, 49*, 814-824.

Ho, K-C., Roessmann, U., Straumfjord, J.V., & Monroe, G. (1980a). Analysis of brain weight: I. Adult brain weight in relation to sex, race, and age. *Archives of Pathology and Laboratory Medicine, 104*, 635-639.

Ho, K-C., Roessmann, U., Straumfjord, J.V., & Monroe, G. (1980b). Analysis of brain weight: II. Adult brain weight in relation to body height, weight, and surface area. *Archives of Pathology and Laboratory Medicine, 104*, 640-645.

Ho, K-C., Roessmann, U., Hause, L., & Monroe, G. (1981). Newborn brain weight in relation to maturity, sex, and race. *Annals of Neurology, 10*, 243-246.

Hofman, M.A. (1991). The fractal geometry of convoluted brains. *Journal fur Hirnforschung, 32*, 103-111.

Hofman, M.A. (1993). Encephalization and the evolution of longevity in mammals. *Journal of Evolutionary Biology, 6*, 209-227.

Hofman, M.A., & Swaab, D.F. (1991). Sexual dimorphism of the human brain: Myth and reality. *Experimental Clinical Endocrinology, 98*, 161-170.

Holloway, R.L. (1980). Within-species brain-body weight variability: A reexamination of the Danish data and other primate species. *American Journal of Physical Anthropology, 53*, 109-121.

Howells, W.W. (1973). *Cranial variation in man*. (Papers of the Peabody Museum of Archaeology and Ethnology, volume 67). Cambridge, MA: Harvard University Press.

Jackson, D.N. (1993, July). Sex differences in intellectual ability. Presented at the International Society for the Study of Individual Differences, Baltimore, Maryland.

Jackson, D.N. (1995, May). Sex differences in general cognitive ability. Presented at the International Behavioral Development Symposium: Biological Basis of Sexual Orientation and Sex-Typical Behavior, Minot State University, Minot, North Dakota.

Jensen, A.R. (1973). *Educability and group differences*. London: Methuen.

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

Jensen, A.R. (1993). Spearman's hypothesis tested with chronometric information processing tasks. *Intelligence, 17*, 47-77.

Jensen, A.R. (1994). Psychometric *g* related to differences in head size. *Personality and Individual Differences, 17*, 597-606.

Jensen, A.R. (1997). Psychometric *g* and the race question. In W. Tomic & J. Kingma (Eds.), Advances in cognition and educational practice (Vol. 4, pp. 1-23). Greenwich, CT: JAI Press.

Jensen, A.R., & Inouye, A.R. (1980). Level I and Level II abilities in Asian, white and black children. *Intelligence, 4*, 41-49.

Jensen, A.R., & Johnson, F.W. (1994). Race and sex differences in head size and IQ. *Intelligence, 18*, 309-333.

Jensen, A.R., & Reynolds, C.R. (1982). Race, social class and ability patterns on the WISC-R. *Personality and Individual Differences, 3*, 423-438.

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* (pp. 139-242). Norwood, NJ: Ablex.

Jensen, A.R., & Whang, P.A. (1993). Reaction times and intelligence: A comparison of Chinese-American and Anglo- American children. *Journal of Biosocial Science, 25*, 397-410.

Jensen, A.R., & Whang, P.A. (1994). Speed of accessing arithmetic facts in long-term memory: A comparison of Chinese-American and Anglo-American children. *Contemporary Educational Psychology, 19*, 1-12.

Jerison, H.J. (1963). Interpreting the evolution of the brain. *Human Biology, 35*, 263-291.

Jerison, H.J. (1973). *Evolution of the brain and intelligence*. New York: Academic.

Kandel, E.R. (1991). Nerve cells and behavior. In E.R. Kandel, J.H. Swartz, & T.M. Jessell (Eds.), *Principles of neural selection* (3rd ed.). New York: Elsevier.

Kimura, D. (1992). Sex differences in the brain. *Scientific American, 267*(3), 119-125.

Klein, R.E., Freeman, H.E., Kagan, J., Yarborough, C., & Habicht, J.P. (1972). Is big smart? The relation of growth to cognition. *Journal of Health and Social Behavior, 13*, 219-250.

Kline, C.L., & Lee, N. (1972). A transcultural study of dyslexia: analysis of language disabilties in 277 Chinese children simultaneously learning to read and write in English and Chinese. *Journal of Special Education, 6*, 9-26.

Kolakowski, D., & Malina, R.M. (1974). Spatial ability, throwing accuracy, and man's hunting heritage. *Nature, 251*, 410-412.

Krogman, W.M. (1970). Growth of head, face, trunk and limbs in Philadelphia White and Negro children of elementary and high school age. *Monographs of the Society for Research in Child Development, 35*, 136.

Lee, A., & Pearson, K. (1901). Data for the problem of evolution in man. VI. A first study of the correlation of the human skull. *Philosophical Transactions of the Royal Society of London, 196A*, 225-264.

Lesser, G.S., Fifer, F., & Clark, H. (1965). Mental abilitites of children from different social class and cultural groups. *Monographs of the Society for Research in Child Development, 30*, 102.

Lewontin, R.C., Rose, S., & Kamin, L.J. (1984). *Not in our genes*. New York: Pantheon.

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

Lynn, R. (1977a). The intelligence of the Chinese and Malays in Singapore. *Mankind Quarterly, 18*, 125-128.

Lynn, R. (1977b). The intelligence of the Japanese. *Bulletin of the British Psychological Society, 30*, 69-72.

Lynn, R. (1990). New evidence on brain size and intelligence: A comment on Rushton and Cain and Vanderwolf. *Personality and Individual Differences, 11*, 795-797.

Lynn, R. (1991a). Intelligence in China. *Social Behavior and Personality, 19*, 1-4.

Lynn, R. (1991b). Race differences in intelligence: A global perspective. *Mankind Quarterly, 31*, 255-296.

Lynn, R. (1993). Further evidence for the existence of race and sex differences in cranial capacity. *Social Behavior and Personality, 21*, 89-92.

Lynn, R. (1994a). The intelligence of Ethiopian immigrant and Israeli adolescents: A comment on Kaniel and Fisherman. *International Journal of Psychology, 29*, 55-56.

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

Lynn, R., & Hampson, S. (1986a). Further evidence on the cognitive abilities of the Japanese: data from the WPPSI. *International Journal of Behavioral Development, 10*, 23-36.

Lynn, R., & Hampson, S. (1986b). Intellectual abilities of Japanese children: an assessment of 2 1/2-8 1/2 year olds derived from the McCarthy Scales of Children's Abilities. *Intelligence, 10*, 41-58.

Lynn, R., & Hampson, S. (1986c). The structure of Japanese abilities: an analysis in terms of the hierarchical model of intelligence. *Current Psychological Research and Reviews, 4*, 309-322.

Lynn, R., Hampson, S., & Bingham, R. (1987). Japanese, British and American adolescents compared for Spearman's g and for the verbal, numerical and visuo-spatial abilities. *Psychologia, 30*, 137-144.

Lynn, R., Hampson, S.L., & Iwawaki, S. (1987). Abstract reasoning and spatial abilities among American, British and Japanese adolescents. *Mankind Quarterly, 27*, 397-434.

Lynn, R., Hampson, S., & Lee, M. (1988). The intelligence of Chinese children in Hong Kong. *Social Psychology International, 9*, 29-32.

Lynn, R., & Hattori, K. (1990). The heritability of intelligence in Japan. *Behavior Genetics, 20*, 545-546.

Lynn, R., & Holmshaw, M. (1990). Black-white differences in reaction times and intelligence. *Social Behavior and Personality, 18*, 299-308.

Lynn, R., & Jindal, S. (1993). Positive correlations between brain size and intelligence: Further evidence from India. *Mankind Quarterly, 34*, 109-123.

Lynn, R., Pagliari, C., & Chan, J. (1988). Intelligence in Hong Kong measured for Spearman's g and the visuospatial and verbal primaries. *Intelligence, 12*, 423-433.

Lynn, R., & Shigehisa, T. (1991). Reaction times and intelligence: A comparison of Japanese and British children. *Journal of Biosocial Science, 23*, 409-416.

Mackintosh, N.J., & Mascie-Taylor, C.G.N. (1985). The IQ question. In *Education for all* [*The Swann Report*]. London: HMSO.

Manley, D.R. (1963). Mental ability in Jamaica. *Social and Economic Studies, 12*, 51-77.

Marshall, J. (1892). On the relations between the weight of the brain and its parts, and the stature and mass of the body, in man. *Journal of Anatomy and Physiology, 26*, 445-500.

Michael, J.A. (1988). A new look at Morton's craniological research. *Current Anthropology, 29*, 349-354.

Miele, F. (1979). Cultural bias in the WISC. *Intelligence, 3*, 149-164.

Miller, E.M. (1994). Intelligence and brain myelination: A hypothesis. *Personality and Individual Differences, 17*, 803-832.

Misawa, G., Motegi, M., Fujita, K., & Hattori, K. (1984). A comparative study of intellectual abilities of Japanese and American children on the Columbia Mental Maturity Scale (CMMS). *Personality and Individual Differences, 5*, 173-181.

Montagu, M.F.A. (1960). *An introduction to physical anthropology* (3rd ed.). Springfield, IL: Charles C. Thomas.

Montie, J.E., & Fagan, J.F. (1988). Racial differences in IQ: item analysis of the Stanford-Binet at 3 years. *Intelligence, 12*, 315-332.

Morton, S.G. (1849). Observations on the size of the brain in various races and families of man. *Proceedings of the Academy of Natural Sciences Philadelphia, 4*, 221-224.

Murdock, J., & Sullivan, L.R. (1923). A contribution to the study of mental and physical measurements in normal school children. *American Physical Education Review, 28*, 209-330.

Notcutt, B. (1950). The measurement of Zulu intelligence. *Journal of Social Research, 1*, 195-206.

Ombredane, A., Robaye, F., & Robaye, E. (1952). Analyse des ŕsultats d'une application experimentale du matrix 38 e 485 noirs Baluba. *Bulletin Contre D'Ätudes et Reserches Psychotechniques, 7*, 235-255.

Osborne, R.T. (1980). *Twins: Black and white*. Athens, GA.: Foundation for Human Understanding.

Osborne, R.T. (1992). Cranial capacity and IQ. *Mankind Quarterly, 32*, 275-280.

Owen, K. (1989). *Test and item bias: The suitability of the Junior Aptitude Tests as a common test battery for White, Indian and Black pupils in standard 7*. Pretoria, South Africa: Human Science Research Council.

Owen, K. (1992). The suitability of Raven's Standard Progressive Matrices for various groups in South Africa. *Personality and Individual Differences, 13*, 149-159.

Packard, G.C., & Boardman, T.J. (1988). The misuse of ratios, indices, and percentages in ecophysiological research. *Physiological Zoology, 61*, 1-9.

Pagel, M.D. & Harvey, P.H. (1988). How mammals produce large-brained offspring. *Evolution, 42*, 948-957.

Pakkenberg, H., & Voigt, J. (1964). Brain weight of the Danes. *Acta Anatomica, 56*, 297-307.

Papiernik, E., Cohen, H., Richard, A., de Oca, M.M., & Feingold, J. (1986). Ethnic differences in duration of pregnancy. *Annals of Human Biology, 13*, 259-265.

Passingham, R.E. (1979). Brain size and intelligence in man. *Brain, Behavior and Evolution, 16*, 253-270.

Pearl, R. (1906). On the correlation between intelligence and the size of the head. *Journal of Comparative Neurology and Psychology, 16*, 189-199.

Pearl, R. (1934). The weight of the Negro brain. *Science, 80*, 431-434.

Pearlson, G.D., Kim, W.S., Kubos, K.L., Moberg, P.J., Jayaram, G., Bascom, M.J., Chase, G.A., Goldfinger, A.D., & Tune, L.E. (1989). Ventricle-brain ratio, computed tomographic density, and brain area in 50 schizophrenics. *Archives of General Psychiatry, 46*, 690-697.

Pearson, K. (1906). On the relationship of intelligence to size and shape of head, and to other physical and mental characters. *Biometrika, 5*, 105-146.

Pedersen, N.L., Plomin, R., Nesselroade, J.R., & McClearn, G.D. (1992). A quantitative genetic analysis of cognitive abilities during the second half of the life span. *Psychological Science, 3*, 346-353.

Piven, J., Arndt, S., Bailey, J., Havercamp, S., Andreasen, N.C., & Palmer, P. (1995). An MRI study of brain size in autism. *American Journal of Psychiatry 152*, 1145-1149.

Pons, A.L. (1974). Administration of tests outside the cultures of their origin. Presented at the 26th Congress of the South African Psychological Association.

Pool, R. (1994). *Eve's rib: Searching for the biological roots of sex differences*. New York: Crown.

Porteus, S.D. (1937). *Primitive intelligence and environment*. New York: Macmillan.

Raz, N., Torres, I.J., Spencer, W.D., Millman, D., Baertschi, J.C., & Sarpel, G. (1993). Neuroanatomical correlates of age-sensitive and age-invariant cognitive abilities: An in vivo MRI investigation. *Intelligence, 17*, 407-422.

Reed, T.E., & Jensen, A.R. (1993). Cranial capacity: New Caucasian data and comments on Rushton's claimed Mongoloid-Caucasoid brain size differences. *Intelligence, 17*, 423-431.

Reid, R.W., & Mulligan, J.H. (1923). Relation of cranial capacity to intelligence. *Journal of the Royal Anthropological Institute, 53*, 322-331.

Ricklan, D.E., & Tobias, P.V. (1986). Unusually low sexual dimorphism of endocranial capacity in Zulu cranial series. *American Journal of Physical Anthropology, 71*, 285-293.

Rodd, W.G. (1959). A cross cultural study of Taiwan's Schools. *Journal of Social Psychology, 50*, 3-36.

Rushton, J.P. (1988a). Race differences in behaviour: A review and evolutionary analysis. *Personality and Individual Differences, 9*, 1009-1024.

Rushton, J.P. (1988b). The reality of racial differences: A rejoinder with new evidence. *Personality and Individual Differences, 9*, 1035-1040.

Rushton, J.P. (1989a). The evolution of racial differences. *Journal of Research in Personality, 23*, 7-20.

Rushton, J.P. (1989b). Japanese inbreeding depression scores: Predictors of cognitive differences between blacks and whites. *Intelligence, 13*, 43-51.

Rushton, J.P. (1990). Race, brain size and intelligence: A rejoinder to Cain and Vanderwolf. *Personality and Individual Differences, 11*, 785-794.

Rushton, J.P. (1991). Mongoloid-Caucasoid differences in brain size from military samples. *Intelligence, 15*, 351-359.

Rushton, J.P. (1992a). Cranial capacity related to sex, rank, and race in a stratified random sample of 6,325 U.S. military personnel. *Intelligence, 16*, 401-413.

Rushton, J.P. (1992b). Life history comparisons between Orientals and Whites at a Canadian university. *Personality and Individual Differences, 13*, 439-442.

Rushton, J.P. (1993). Corrections to a paper on race and sex differences in brain size and intelligence. *Personality and Individual Differences, 15*, 229-231.

Rushton, J.P. (1994). Sex and race differences in cranial capacity from International Labour Office data. *Intelligence, 19*, 281-294.

Rushton, J.P. (1995). *Race, evolution and behavior: A life-history perspective*. New Brunswick, NJ: Transaction Publishers.

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

Rushton, J.P., & Osborne, R.T. (1995). Genetic and environmental contributions to cranial capacity estimated in Black and White adolescents. *Intelligence, 20*, 1-13.

Scarr, S. (1987). Three cheers for behavior genetics: Winning the war and losing our identity. *Behavior Genetics, 17*, 219-228.

Scarr, S., Caparulo, B.K., Ferdman, B.M., Tower, R.B., & Caplan, J. (1983). Developmental status and school achievements of minority and non-minority children from birth to 18 years in a British Midlands town. *British Journal of Developmental Psychology, 1*, 31-48.

Scarr, S., Weinberg, R.A., & Waldman, I.D. (1993). IQ correlations in transracial adoptive families. *Intelligence, 17*, 541-555.

Schreider, E. (1968). Quelques corrälations somatiques des tests mentaux. *Homo, 19*, 38-43.

Schultz, R.T. (1991). The relationship between intelligence and gray-white matter image contrast: A MRI study of healthy college students. Unpublished doctoral dissertation, University of Texas, Austin.

Shibata, I. (1936). Brain weight of the Korean. *American Journal of Physical Anthropology, 22*, 27-35.

Shigehisa, T., & Lynn, R. (1991). Reaction times and intelligence in Japanese children. *International Journal of Psychology, 26*, 195-202.

Shuey, A.M. (1966). *The testing of Negro intelligence*. New York: Social Science Press.

Simmons, K. (1942). Cranial capacities by both plastic and water techniques with cranial linear measurements of the Reserve Collection; white and Negro. *Human Biology, 14*, 473-498.

Sinha, U. (1968). The use of Raven's Progressive Matrices in India. *Indian Educational Review, 3*, 75-88.

Sommerville, R.C. (1924). Physical, motor and sensory traits. *Archives of Psychology, 12*, 1-108.

Sorokin, P. (1927). *Social mobility*. New York: Harper.

Spitzka, E.A. (1903). The brain-weight of the Japanese. *Science, 18*, 371-373.

Steen, L.A. (1987). Mathematics education: A predictor of scientific competitiveness. *Science, 237*, 251-253.

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

Stumpf, H., & Jackson, D.N. (1994). Gender-related differences in cognitive abilities: Evidence from a medical school admissions testing program. *Personality and Individual Differences, 17*, 335-344.

Susanne, C. (1979). On the relationship between psychometric and anthropometric traits. *American Journal of Physical Anthropology, 51*, 421-423.

Swaab, D.F., & Hofman, M.A. (1984). Sexual differentiation of the human brain: A historical perspective, In G.J. De Vries et al. (Eds.), *Progress in brain research* (Vol. 61, pp. 361-374). Amsterdam: Elsevier.

Takahashi, K., & Suzuki, I. (1961). On the brain weight of recent Japanese. *Sapporo Medical Journal, 20*, 179-184.

Terman, L.M. (1926/1959). *Genetic studies of genius: vol. 1. Mental and physical traits of a thousand gifted children* (2nd ed.). Stanford, CA: Stanford University Press.

Tobias, P.V. (1970). Brain-size, grey matter and race—fact or fiction? *American Journal of Physical Anthropology, 32*, 3-26.

Todd, T.W. (1923). Cranial capacity and linear dimensions in white and Negro. *American Journal of Physical Anthropology, 6*, 97-194.

Vernon, P.E. (1969). *Intelligence and cultural environment*. London: Methuen.

Vernon, P.E. (1982). *The abilities and achievements of Orientals in north America*. New York: Academic.

Vint, F.W. (1934). The brain of the Kenya native. *Journal of Anatomy, 48*, 216-223.

Voigt, J., & Pakkenberg, H. (1983). Brain weight of Danish children. *Acta Anatomica, 116*, 290-301.

Voyer, D., Voyer, S., & Bryden, M.P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*, 250-270.

Wainer, H. (1988). How accurately can we assess changes in minority performance on the SAT? *American Psychologist, 43*, 774-778.

Weinberg, W.A., Dietz, S.G., Penick, E.C., & McAlister, W.H. (1974). Intelligence, reading achievement, physical size and social class. *Journal of Pediatrics, 85*, 482-489.

Weinberg, R.A., Scarr, S., & Waldman, I.D. (1992). The Minnesota Transracial Adoption Study: A follow-up of IQ test performance at adolescence, *Intelligence, 16*, 117-135.

Wickett, J.C., Vernon, P.A., & Lee, D.H. (1994). In vivo brain size, head perimeter, and intelligence in a sample of healthy adult females. *Personality and Individual Differences, 16*, 831-838.

Willerman, L., Schultz, R., Rutledge, J.N., & Bigler, E.D. (1991). In vivo brain size and intelligence. *Intelligence, 15*, 223-228.

Winick, M., Meyer, K.K., & Harris, R.C. (1975). Malnutrition and environmental enrichment by early adoption. *Science, 190*, 1173-1175.

Witelson, S.F., Glezer, I.I., & Kigar, D.L. (1995). Women have greater density of neurons in posterior temporal cortex. *Journal of Neuroscience, 15,* 3418-3428.

Wober, M. (1969). The meaning and stability of Raven's Matrices Test among Africans. *International Journal of Psychology, 4*, 229-235.

Yeo, R.A., Turkheimer, E., Raz, N., & Bigler, E.D. (1987). Volumetric asymmetries of the human brain: Intellectual correlates. *Brain and Cognition, 6*, 15-23.

Zindi, F. (1994). Differences in psychometric performance. *The Psychologist, 7*, 549-552

ON THE COMPARATIVE GENERALITY OF *g*

Charles Locurto

INTRODUCTION

As a way of introducing the problem of the comparative generality of *g* or general intelligence, consider the following quotes. In particular, compare the spirit and tone of the first two quotes with the latter two:

> We interpret the preponderance of the evidence as overwhelmingly supporting the existence of some kind of general factor in human intelligence. Indeed, we are unable to find any convincing evidence at all that militates against this view (Sternberg & Gardner, 1982, p. 231).
>
> Spearman's *g* is surely one of the most interesting and enduring constructs in all of psychology (Jensen, 1985, p. 194).
>
> There are no within-species differences (that is, individual differences) in intellect in nonhuman vertebrates, and indeed there is no evidence for nonhuman individual differences in general intelligence (Macphail, 1985b, p. 234).
>
> One cannot speak of (general) intelligence within a species of animals. No one has found evidence in support of a general level or capacity that result (sic) in an animal ... consistently performing above or below the level attained by other members of the species on several different tasks (Warren, 1977, pp. 41-42; parenthetical added).

Advances in Cognition and Educational Practice, Volume 4, pages 79-100.

ISBN: 0-7623-0105-8

The first two quotes evidence a tone that is profoundly positive, suggesting as they do the existence of a construct with the power to qualify as one of psychology's rare universals. The latter two quotes, though they reveal almost the same level of certainty, are decidedly negative with respect to the existence of a general factor in intelligence. The difference between these two sets of quotes is that the first two are referring to a large body of evidence regarding human intelligence, whereas the latter two are referring to what will be shown to be a much smaller array of evidence regarding animal intelligence. It is this difference and its interpretation that constitute the focus of this chapter.

GENERAL INTELLIGENCE IN HUMANS

The Evidence For Human g

The idea that a general intellectual ability exists in human behavior comes from the sort of data presented in Table 1. The table displays the results of administering a typical intelligence battery, the Multidimensional Aptitude Battery (MAB), to 170 undergraduates (adapted from Miller & Vernon, 1992, Table 3, p. 18). The most salient characteristic of the data presented in Table 1 is that the 10 tasks composing the battery correlate positively with each other.

The average correlation across this matrix is .29 (range = .06-.62). This pattern of correlations, a positive manifold as it is called, is not at all unusual. Beginning with Charles Spearman who first reported its existence in 1904, this finding is a common outcome of administering a battery of complex intellectual tasks to human subjects. It is this positive manifold that has given rise to the notion that apart from the unique variance associated with each task (which Spearman referred to as **s** or specific variance), some of the variance in this matrix is common across all tasks. Stated differently, the relative stability of individual differences across the different tasks suggests that some sort of general ability is involved in the processing and solution of every task in addition to the specific mechanisms evoked by that task.

The magnitude of this general ability can be seen when the matrix of positive correlations is subjected to factor analysis. A wide variety of analyses have demonstrated that all subtests load significantly on the first factor. For the data in Table 1, for example, the first factor accounted for 37 percent of matrix variance, and subtest loadings were all positive, ranging from .36 to .78. The first factor not only accounts for more of the total matrix variance than any other factor, but it also often accounts for more variance than any *combination* of other factors (Jensen, 1986; Thorndike, 1994). In the MAB, for example, the ten subtests can be grouped into two lower-order factors, verbal (the first five subtests) and performance (the second five subtests). These two lower-order factors together accounted for less than 20 percent

Table 1. Correlation Matrix for Multimensional Aptitude Battery

Test	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Information (1)	1.00									
Comprehension (2)	0.53	1.00								
Arithmetic (3)	0.29	0.51	1.00							
Similarities (4)	0.59	0.45	0.22	1.00						
Vocabulary (5)	0.62	0.06	0.16	0.57	1.00					
Digit Symbol (6)	0.12	0.25	0.12	0.20	0.14	1.00				
Picture Comp. (7)	0.48	0.15	0.10	0.33	0.33	0.14	1.00			
Spatial (8)	0.18	0.19	0.42	0.17	0.14	0.30	0.33	1.00		
Picture Arrange. (9)	0.39	0.07	0.24	0.39	0.31	0.26	0.44	0.23	1.00	
Object Assembly (10)	0.20	0.69	0.41	0.23	0.08	0.24	0.32	0.52	0.33	1.00

of matrix variance. A similar example comes from the latest revision of the Wechsler Preschool and Primary Scale of Intelligence-Revised (WPSSI-R), a widely used intelligence test applicable to children in the age range of three to seven years. Factor analysis of the standardization sample ($n = 1800$) for the WPSSI-R indicated that the first factor accounted for approximately 33 percent of the total variance across the different age groupings. In contrast, the two lower-order or group factors extracted in this analysis, verbal/educational and spatial/perceptual factors (analogous to the lower-order verbal and performance factors of the MAB), together accounted for only 16 percent of total variance (Blaha & Walbrown, 1991).

Alternative Interpretations of the Positive Manifold

To be sure, there is almost always more than one possible factor analysis and interpretation of a correlation matrix, and, historically, there have been a number of attempts to bypass the powerful interpretation of a general intellectual ability. Perhaps the most radical contemporary alternative has been to dismiss the idea of *g* altogether, an approach advanced most conspicuously by Howard Gardner (1983) in his theory of "multiple intelligences" (see also, Gould, 1981, for a similar position). Gardner proposes a theory of the sort that Spearman (1927) years earlier had referred to as "anarchic" in form. Human intellect, argues Gardner, consists of seven independent types of abilities: linguistic, spatial, logical/mathematical, bodily/kinesthetic, musical, intrapersonal and interpersonal/social. The core of his approach is to suggest that an individual may rank quite high or low on one of these intelligences and be at the other end of the continuum on another. In short, Gardner argues that these different abilities exist without the cohesiveness of *g* to integrate them.

On the face of it, Gardner's theory would seem to be limited by the fact that several of his "intelligences" such as linguistic, spatial, and logical/mathematical constitute the heart of standard intelligence tests. There is no

doubt that these abilities share significant common variance. How, then, can they be referred to as independent? Gardner's (1983) answer is essentially that all subtests in a standard intelligence battery are likely to be mediated by what he calls "linguistic and logical facility" (p. xi), regardless of whatever else they measure. As a result, their apparent linkage may be artificial to some degree, and although a general factor may be derived from standard batteries of intelligence tests it is of limited value because it may fail to identify individuals with particular capability in other sorts of intelligences.

Unfortunately, Gardner does not buttress his ideas with quantitative analyses of the sort that have become standardfare in intelligence testing. His database is instead composed principally of case studies and qualitative examples drawn from anthropology and cross/cultural investigations. As a result, there is no evidence presented that his supposed separate intelligences are indeed independent, or that they are to a large degree not captured by the *g* derived from the factor analysis of standard test batteries.

Additionally, if Gardner is correct that *g* is saturated with "linguistic and logical facility," that fact alone may be taken to indicate not the artificial nature of standard intelligence tests or the irrelevance of *g*, but rather their central importance, at least in Western cultures that undoubtedly value just those abilities. Surely, it might be argued, success in modern cultures requires a modicum of linguistic and logical competence, whatever else might constitute the nature of an individual's particular array of talents or "intelligences" (see also Sternberg, 1986, for a similar point). The issue here might be recast by asking whether the *g* extracted by factor analysis has predictive value. To anticipate a later section of this chapter, no one except diehard psychometricans should be interested in *g* unless it can be demonstrated that *g* tells us something about the world beyond the factor matrix. Does it predict, for example, the sorts of educational and societal outcomes that we might expect of a true measure of intelligence? If it does not, then *g* may indeed be of little extended value. If it does so predict, then perhaps Gardner's dismissal of standard tests, and of *g*, is mistaken.

As a further point, suppose we view Gardner's arguments about different sorts of intelligences in a somewhat weaker form, as suggesting limits to the importance of *g* but not its outright abandonment by the scientific community. After all, no single measure can account for all that we mean commonsensably by the word intelligence. This point is undoubtedly true; the problem with it is only that it is not at all original. Consider, for example, the following statement taken from a description of the inaugural publication of the first widely used intelligence test in the United States, the Stanford-Binet, in 1916:

> The scale does not pretend to bring to light the idiosyncrasies of special talent, but only to measure the general level of intelligence. It cannot be used for the discovery of exceptional ability in drawing, painting, music, mathematics, oratory, salesmanship, etc. because no

> effort is made to explore the processes underlying these abilities. It can, therefore, never serve as a *detailed chart* for the vocational guidance of children, telling us which will succeed in business, which in art, which in medicine, etc. (Terman, 1916, p. 49; italics in original).

A most curious feature of this statement is not simply that it anticipated Gardner's problems with the limited predictive validity of intelligence tests. It may be surprising to some that the statement was written by Lewis Terman who is regularly vilified as one of the evil hereditarians of the early history of testing in the United States, someone for whom the IQ (Intelligence Quotient) was destiny (e.g., Gould, 1981; Kamin, 1973; see also Terman & Merrill, 1937, for a continuation of this theme in the next revision of the Stanford-Binet). This statement indicates fully that the limitations of intelligence tests, and, concomitantly of *g*, and the existence of important human abilities unmeasured by *g*, have been well-known since the inception of intelligence testing.[1]

It should also be noted that the existence of abilities not fully accounted for by *g* was extensively studied in the 1930s and 1940s by Thurstone (e.g., Thurstone, 1947). Using a form of factor analysis that he termed multiple factors, Thurstone identified eight "primary abilities" including verbal, number, spatial, word fluency, deductive reasoning, inductive reasoning, rote memory, and perceptual speed. His analysis, based initially on a battery of 56 different tests given to 240 college students, seemed to indicate a complete revamping of Spearman's *g*. Thurstone reported that these primary abilities were not linked to a general factor but were distinct entities. Thurstone's work is often presented as an inhospitable challenge to the centrality of *g* in psychological theory (see Gould, 1981).

The historical record does not support that conclusion, however. Immediately following the initial publication of Thurstone's position the central problem in his analysis was identified by Spearman and others, namely that the primary abilities were themselves correlated. These correlations of course suggested that Thurstone's primaries could themselves be factored to reveal *g*. A subsequent set of re-analyses of Thurstone's data revealed just this conclusion. In one such study *g* accounted for 31 percent of total matrix variance in Thurstone's data, whereas the combination of all the primaries accounted for just 24 percent (Vernon, 1965). Other similar analyses reported by Thorndike (1994) indicated that the *g* extracted from Thurstone's primary abilities accounted for nearly 50 percent of total variance. Ironically perhaps, Thorndike also reports that a commercially-developed test employing Thurstone's primaries has consistently been used by test administrators to derive a single overall score of intellectual development.

Additionally, soon after Thurstone's publication of his primary abilities Raymond Cattell (1987) reports that he suggested to Thurstone that *g* and the primary mental abilities could be reconciled:

> (I) pointed out, and Thurstone fully concurred, that since a "g" factor could be obtained as a *second-order* factor among his primaries, the concept of *g* was still firm. The Spearman and Thurstone findings were reconcilable; and with mutual illumination ... Thus he (Thurstone) was under no illusion that he had succeeded in abolishing *g*. Indeed, he pointed out an extremely important though subtle sense in which his approach to "g" through primaries actually defined "g" in a more scientifically satisfactory way than it had been defined by Spearman (p. 30; parentheticals added).

Subsequent work further illuminated the link between the specific tasks that form an intelligence battery, Thurstone's primary abilities, and *g*, in the form of a hierarchical theory of intelligence (Vernon, 1965). Figure 1 presents a rather complex elaboration of this theory as a three-stage structure that has *g* at its apex. In this sort of model a large number of individual tests are first reduced to a set of primary factors such as verbal fluency, verbal comprehension, reasoning, numerical operations, memory, and the like. At a higher level, these factors, owing to their correlations, can be represented by two second-order factors such as the verbal and performance factors of the MAB. As these factors are themselves correlated, they reveal a general factor as a third-order factor.[2]

The model presented in Figure 1 is but one a number of possible hierarchical configurations (see Gustafsson, 1988, for the explication of a number of other hierarchical models of cognitive abilities). What is common to any such model is the increasing breadth of higher-level factors and the existence of a general factor at the apex of the structure representing a component of variance that

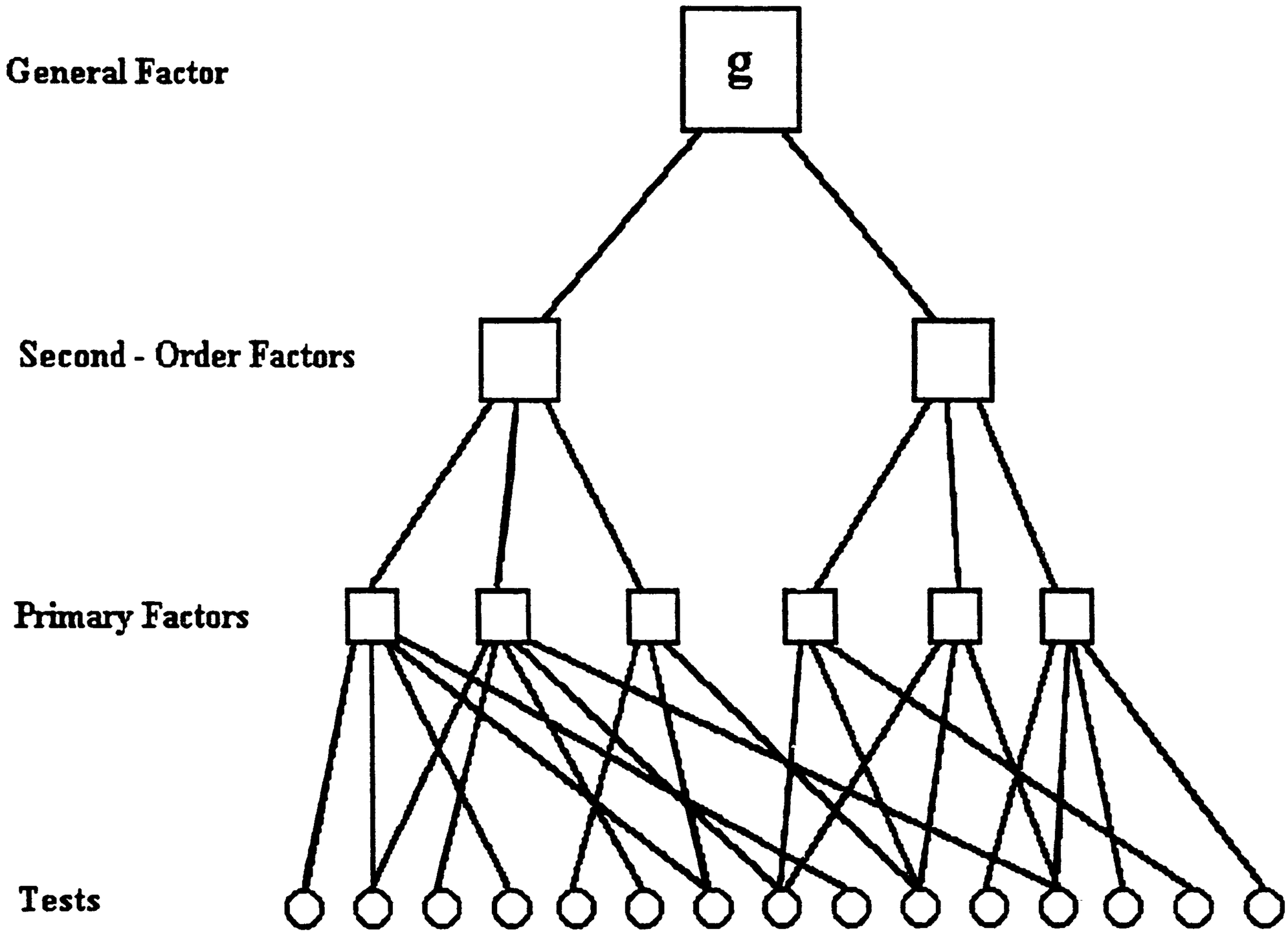


Figure 1. Hierarchical model of Human Intelligence

all the specific tasks have in common. As such, the existence of *g* is not antagonistic to Thurstone's primaries but appears instead as a powerful organizing factor for all lower-order factors, and one that cannot easily be described in terms of any specific information content or set of skills, as might be true for lower-order factors.

What this model, or any other model, cannot tell us, however, is whether there is any reality to this general factor beyond the set of tasks upon which it is based. As stated earlier, the issue is whether *g* is simply the inevitable outcome of factoring any matrix of human intellectual tasks that are in some way verbally mediated, or whether *g* can be demonstrated to have some reality beyond the factor matrix.

Life Beyond the Factor Matrix: The Nature of *g*

A first approach to evaluating *g*'s status beyond factor analysis is to ask whether the *g*s derived from different test batteries are themselves intercorrelated, suggesting at least that different batteries are measuring the same *g*. The answer appears to be yes (Jensen, 1986). Although *g* is not entirely invariant across different tests (Brody, 1992), its stability across widely differing test batteries is impressively high, averaging better than 0.83 in one extended study of differently-constructed batteries (Thorndike, 1987, 1994). The same *g*, it seems, is likely to reveal itself in almost any group of tests. In addition, as test batteries grow larger, so do the correlations between their *g* measures, suggesting that larger batteries are more accurately sampling the same underlying construct. The *g* derived from different tests is also rather stable across different types of measurement strategies, so much so that almost any method of determining *g* appears to be as good (i.e., reveals the same magnitude *g*) as any other (Jensen & Weng, 1994; Ree & Earles, 1991). As a final point, the *g* extracted from different ethnic groups appears to be the same. A recent study of the Wechsler Intelligence Scale for Children-Revised (WISC-R) given to white, black, and Hispanic children aged ($n = 560$) 6 to 11 years indicated that both the magnitude and pattern of factor loading on all subtests were virtually identical across these three ethnic groups (Taylor & Ziegler, 1987).

These findings coalesce in the direction of supporting the robust nature of *g*. Yet, reasonably, it might be objected that they are all psychometric in nature; many of them are findings internal to the test batteries themselves. A true general ability would be expected to reveal itself in criteria that are nonpsychometric in nature. In recognition of the importance of this problem, an extensive body of evidence has been developed that identifies criteria external to the batteries themselves to demonstrate the reality of *g*. It has long been known, for example, that intelligence test scores are substantially correlated with educational achievement and occupational status (e.g., Brody, 1992; Jensen, 1981; Locurto, 1991). What may not be as well known is that

it is *g*—rather than variance associated with lower-order factors—that constitutes the largest component of what Jensen and Weng (1994) refer to as the "practical validity" of intelligence tests in these areas. Conversely, when *g* variance is partialled out of test scores, test validity in these realms reduces almost to nonsignificance (Jensen, 1986).

The relation of *g* to occupational performance is of particular interest. The origins of individual differences in occupational success have often been a contentious issue given the disposition of some investigators to attribute occupational success to non-test-related factors such as motivation and specific job training rather than to differences in intelligence per se. Recent analyses by Gottfredson (1986), however, demonstrate that *g* is extremely important in predicting all aspects of occupational performance (correlations average about 0.50 between *g* and occupational performance ratings, for example). *g* is a useful predictor across a wide range of occupations, but as might be expected, is especially useful in predicting success in more complex occupations. In these sorts of occupations *g* is often more predictive than tests constructed to measure specific types of occupational performance.

From these socially-relevant validities *g* appears to have life beyond the factor matrix. It might still be objected that even these measures are mediated by linguistically/cognitive abilities, but there is also considerable evidence that *g* may be have *biological reality* as well. If one examines the heritability of the various subtests that compose an intelligence battery, for example, that pattern of heritabilities is directly related to a subtest's *g*-loading: the higher the heritability of a given subtest, the higher the *g*-loading of that subtest tends to be (correlations that average about 0.58; see Carden, Fulker, DeFries, & Plomin, 1992; Jensen & Weng, 1994).[3] Other genetically-related effects on subtests and total IQs have been found that also appear to be *g*-mediated. For example, inbreeding depression, the finding that the IQs of offspring of genetically-related individuals are lower than the scores of offspring of unrelated individuals, has been regularly observed in IQ (Bouchard, 1993). The magnitude of this effect on a particular subtest is also directly related to a subtest's *g*-loading (Jensen, 1986).

Of particular interest with regard to the nonpsychometric validity of *g* is the growing evidence that *g* is significantly related to a number of measures that bear no obvious relationship to standard test batteries. These relationships generally sort into two categories. At the behavioral level there is a group of what are referred to as elementary cognitive tasks (ECTs), including simple and discriminated reaction time, and inspection time (the time required to distinguish between two lines of different lengths). These tasks have been shown to correlate substantially with total IQ scores and, as may by now be expected, especially with *g* (Brody, 1992; Eysenck, 1979, 1993; Jensen, 1985; Vernon, 1987). The range of these correlations may average as high as .45 to .55 for inspection time tasks (Nettleback, 1987).[4] A related strategy has examined the

relationship between IQ and measures of infant habituation (e.g., decrements in responding that follow repeated presentations of a stimulus). Habituation measures are of interest because they can be obtained within the first few months of life, years before IQs can be reliably measured. A summary of 14 studies indicated an average correlation of .40 to .50 between habituation measures obtained on average at six months and IQs obtained at 4 to 6 years of age. When corrected for attenuation, these correlations may be as high as .70 to .75 (Brody, 1992, Table 3.2, pp. 74-75; separate correlations with *g*-loadings per se were not presented).

A second group of tasks are physiological in nature. The findings in this second group vary considerably, and the reasons for this variability are as yet not understood. Despite this caution, there is a considerable evidence that measures such as EEGs, averaged evoked potentials (AEPs), and glucose metabolism, correlate reliably with IQ (see Vernon, 1993, for a review). Taken together, the relations between *g* and ECTs, infant habituation, and these physiological measures suggest that *g* may be best understood as a reflection of the speed and/or efficiency of processing information. That possibility is at once intriguing and unifying, although no definitive causal inferences can be drawn at present. It is premature to argue as has Jensen that the fundamental (biological) reality of *g* is so well documented that "g links psychology to biology and evolution" (Jensen, 1985, p. 197).

Yet it can be said even at this juncture that previous criticisms of *g*-theory that have rested upon the failure to establish the reality of *g* beyond the factor matrix are now anachronisms. Gould (1981), for example, based much of his criticism of intelligence tests on the failure to document the biological reality of *g*. Commenting on what he viewed as the inability to distinguish between *g*-theory and any of a number of other (anarchic) approaches to human intelligence, Gould argued that "to vindicate either view, some legitimate appeal must be made outside the abstract mathematics [of factor analysis] itself.... All combatants [have] made appeals to biology and advanced tenuous claims, but no concrete tie has even (sic) been confirmed between any neurological object and a factor axis" (p. 310; parenthetical added). The impressive variety of indices which now appear to mark *g* using physiological/biological criteria indicates not only that Gould's view is decidedly mistaken, but also that it may be possible in the near future to assay *g* using methods that bypass entirely the use of traditional intelligence batteries, and therefore avoid many of the criticisms of standardized "paper-and-pencil" tests.

From the perspective of a comparative analysis of *g*, the curious aspect of these findings is that in revealing so broad a foundation for *g*, one that may indeed be biologically-based, it is all the more surprising that *g* is not found in animals, as the quotes at the beginning of this chapter suggest. Could a construct that appears to be profoundly central to human mental functioning have arisen *de novo* in our species?

GENERAL INTELLIGENCE IN ANIMALS?

It was not long after Spearman had introduced the concept of *g* to the scientific community that a number of animal behaviorists attempted to identify an animal *g*. An example of this early work comes from Tomlin and Stone (1934) who studied 132 rats across three learning tasks, two spatial mazes and one visual discrimination task. They reported substantial correlations for the two maze tasks (averaging .43) but a zero-order correlation between the maze tasks and the visual problem. Their conclusion was that the visual and spatial tasks must be measuring "different learning functions" (p. 87). More to the point, they suggested that these findings along with others "would seem to be a discovery of capital importance for students of individual differences in animal learning because it clearly establishes the fact that two fairly reliable instruments may have no community of function" (p. 73). Tomlin and Stone's findings are not dissimilar to the findings of other studies over the past 60 years. A representative sample of this work is presented in Table 2.[5]

The Historical Context of Animal *g*

To anticipate the discussion that follows concerning the particulars of these studies, the principal issue is whether they were sufficiently well-designed so as to permit a conclusion about animal *g*. As is apparent from Table 2, many of these studies are quite dated. Experimental attention to the problem of animal *g* peaked during the mid-1930s when there were a number of studies like that of Tomlin and Stone which found no significant correlations across spatial and visual-type problems in rodents, and therefore concluded that no animal *g* was evident. By the 1940s it had been generally conceded that animal *g* did not exist, and interest in this problem waned considerably. Unfortunately, the model of a hierarchical structure of human abilities became widely known only in the late 1940s (see Burt, 1949, for a review of the development of hierarchical analyses). Royce, in reviewing the available animal literature on individual difference soon thereafter, in 1950, and again in 1966, noted that in the intervening 16 years there had been only four studies generally concerned with the factor structure of animal behavior—and none of them were concerned with animal intelligence per se, much less with a hierarchical model of animal intelligence.

Royce's findings should be seen in light of the fact that this early work was accomplished against a background of no distinctive heritage of studying individual differences in animals. In surveying the available literature in 1942, Tryon lamented that there had been no substantially original studies on the problem of individual differences in animal intelligence since 1934: "Most psychological study of animals continues to consist of experiments showing the effects of laboratory-induced changes of conditions on the *average* animal.

Table 2. Summary of Studies of Individual Differences in Animal Intelligence

Study	Year	Subjects (Number)	Task Descriptions	Average Correlation
Ligget	1925	chicks (90)	2 mazes	.10 overall
Commins, Mcnemar, and Stone	1932	rats (152)	2 mazes, 1 visual discrimination	.60 between mazes; 03 between mazes and discrimination
Campbell	1932	rats (28)	2 mazes, 1 "reasoning" task	.55 between mazes; -.30 between mazes and reasoning task
Dunlap	1933	chicks (133)	4 mazes	.21
Tomlin and Stone	1934	rats (132)	2 mazes, 1 visual discrimination	.43 between mazes; .02 between mazes and discrimination
Thorndike	1935	rats (64)	10 mazes and activity measures	.18 overall; .45 between mazes
Anastasi, Fuller, Scott, and Schmitt	1958	dogs (73)	17 variables of learning and activity	.13 overall
Warren	1961	cats (21)	discriminations and double alternations	.57 between mazes;.19 between discriminations and alternation tasks
Rajalakshimi and Jeeves	1968	rats (60)	maze and visual discrimination	.32 overall
Livsey	1970	rats (8), rabbits (10), cats (8)	maze, discrimination task, and activity measure	.42 overall
Thompson, Crinella & Yu	1987	rats (75)	3 spatial tasks, 1 visual task	.07 (sham-operated controls)
Anderson	1993	rats (22)	2 spatial tasks, 1 novelty preference	.33 overall

Systematic studies of the psychological nature and causation of individual differences are rare" (p. 358, italics in original). This same state of affairs applies to subsequent work.

There is a sense in which differences between animals have been studied, but the focus has been on differences between species, not individuals (see Hulse, 1993, for a recent summary).[6] One approach to species differences has been to assume that the learning mechanisms evidenced by one species may be qualitatively different or even absent in another. This position derived initially from ethologists and was transposed to experimental psychology by

Bitterman (1965) whose work indicated the presence of qualitatively different approaches to problem-solving between different species. A contemporary form of this position has been termed the "adaptive specialization" of animal intelligence (e.g., Seligman, 1970), and holds that "each instance of learning must be treated as a specialized capability shaped by selective pressures and understandable only by reference to the ecology of the animal or of its ancestors—that there are no general laws of learning at all" (Bitterman, 1975, p. 188). Taken to its extreme this position asserts that each species has developed particular and perhaps unique mechanisms that allow it to negotiate its specific ecological niche. Those mechanisms will be similar across different species only to the extent that ecological demands and selection pressures have been similar. This perspective would be unlikely to posit an animal *g* on the grounds that since adaptation is always situation-specific, general mechanisms would have had no reason, no selection pressure, to have evolved.

A rather different position, but one that leads to the same conclusion regarding animal *g*, can be traced to E. L. Thorndike's development of instrumental conditioning (e.g., Bitterman, 1969; Thorndike, 1911). Thorndike noted that a wide variety of animals (monkeys, cats, dogs, fish, and chickens) were capable of solving the same sorts of problems, albeit perhaps at different rates (see Bitterman, 1965; Mackintosh, 1988). Differences between species appeared, in other words, to be quantitative not qualitative, suggesting that the underlying mechanisms were the same as a corollary, if there are only differences of degree between different species in problem solving, there is no reason to posit systematic differences between individuals within the same species. An extreme form of this approach has been developed by Macphail (1982, 1985a, 1985b, 1987) who argues that any instance of supposed qualitative differences between species can be understood as resulting from differences in motivational or informational system employed, or due to the specifics of the testing environment, that is, due to what have been termed "contextual variables." If changed, these contextual variables may drastically alter the results of comparing different species.

Macphail's position may appear to contradict common sense with respect to species differences (Is not a primate capable of more complex learning than a mouse?) but it is not easily disproved: any instance of species differences can be interpreted as a difference in the relevant contextual variables, with the suggestion that some alteration of the experimental environment might show equivalent learning between the involved species. With respect to animal *g* Macphail has argued, as one of the quotations that begins this chapter illustrates, that "one implication, though not a necessary consequence, of this hypothesis, is that there are no within-species differences (that is, individual differences) in intellect in nonhuman vertebrates" (p. 234).

These historical trends provide no basis, either in terms of empirical or theoretical work on individual differences, for an understanding of animal *g*.

In turn, this history leads naturally to the question of whether the available literature, much of it based on work that was accomplished prior to the development of the predominant model of human *g*, meets what might be called the technical requirements for an analysis of animal *g*.

The Necessities for Observing Animal g

Individual Differences

A common feature of the early studies in Table 2 is that many of them provide little information about their subjects. Contemporary studies routinely report the specific strain employed, their ages at the time of testing, sex, and any other relevant characteristics. Older studies are quite variable in these respects. Campbell (1935), for example, reported only that his subjects were "Thirty-four female rats from four to six months of age" (p. 72). Thorndike (1935) described his subjects as "male albino rats, approximately 60 days of age" (p. 20). Tomlin and Stone were a bit more detailed, describing their subjects as "animals taken from 34 litters of the inbred albino stock of the Stanford psychology laboratory" (p. 78).

Unfortunately, the description provided by Tomlin and Stone points to two serious flaws. First, there is the essential need in correlational work for sufficient range and variability. Given that *g* in human studies appears to be highly heritable, it is also important in animal work to use subjects that have appreciable genetic variability. Inbred strains would present a distinct problem in this regard because they are virtually homozygotic as a result of many generations of sibling inbreeding. Given the current state of knowledge regarding genetic influences on animal behavior it is exceedingly unlikely that a contemporary investigator would report correlations gathered from inbred strains (see Plomin, DeFries, & McClearn, 1990, for discussions of genetic influences on a wide range of animal and human behavior; see, however, Crinella & Yu 1995). In this light, it is curious that positive correlations among mazes were often reported in this early work. This finding indicates a vital point about the structure of animal intelligence and its origins to which we will have reason to return.

Second, it would also be unlikely today that an investigator wishing to study a range of problems, some of which utilize visual information, would use albino strains. There is considerable evidence that albinos suffer from comparatively poor visual ability and there are as well some degenerative aspects associated with albinism (Creel, 1980; Wahlston, 1972). It is unclear how frequently albinos were used in this early work given the often brief subject descriptions, but it is apparent that many of these studies used inbred strains. Interestingly, the problem of restricted range that attends using inbred strains was not entirely unknown to early investigators, though the seriousness

of the problem was perhaps not appreciated given the paucity of data regarding the heritability of human *g* at the time. Commins, McNemar and Stone (1932) cited the problem in the Introduction to their study, reporting that "most experimenters use inbred stock which may be expected to be relatively homogeneous or limited in range of ability" (p. 227). They then went on to report substantial correlations between maze tasks but zero-order correlations between spatial tasks and a visual task using "rats from an inbred colony" (p. 229). They concluded as had Tomlin and Stone that there was no "community of function" (p. 234) between spatial and visually-based problem solving, but they also alluded to the presence of a "general learning ability for maze situations" (p. 234).

Tryon (1942) reviewed much of this early work and also noted that most of the work on general intelligence in animals had at that time used inbred stock (see footnote 4, p. 337). In agreement with Tomlin and Stone, he reported that his own work and that of others on the problem of maze learning vs. other types of learning had pointed to the conclusion that "ability seems to be highly specific: whereas it is consistent and constant in a *given* situation, between quite different situations ... it appears to be highly variable" (p. 360, italics in original). For Tryon "there is no such fact as *general* learning ability, but rather that there are *numerous learning abilities*" (p. 356, italics in original).[7]

This early work and Tryon's conclusions about it set the tone for future investigators to pay scant attention to the problem of individual differences and animal *g*. Unfortunately, based simply on the types of subjects used, it appears that Tryon's conclusions are at the least premature. Subject variance is not, however, the most important limitation of these early studies, or those that followed.

Task Variety to "Mark" a Factor

For the construction of a test battery to explore factor structures the most important requirement is the use of a variety of tasks. In the case of animals, task variety might entail the sampling of different information modalities (spatial, visual, auditory, and olfactory in the case of rodents) and behavioral domains (positively and aversively motivated behaviors, for example). Each such modality or domain, in turn, should be sampled by more than one task. As Vernon (1965) noted in referring to hierarchical factor structures, "In general a minimum of three tests is needed to define a factor, hence, only a few factors can be resolved in any one investigation ... if [a] battery consists only, or predominantly, of a specialized type of test, the *g* and major group factors may fail to reveal themselves" (p. 26; parentheticals in original).

In short, no one test can be taken to reveal the presence of a factor since any single test loads only partially on a given factor; the remaining variance may be related to other factors, specific variance associated only with the particular

task, and/or measurement error. Similarly, in the context of a hierarchical model the emergence of *g* itself depends on the measurement of a number of factors using several different tests to mark each lower-order factor. Any two tests may correlate only modestly, depending on their particular pattern of loadings. It is even possible that as Child (1990) noted, "in broad, heterogeneous factors [of which *g* is certainly one], two variables [may] belong to the same factor, and yet have little or no correlation" (p. 33; first parenthetical added). Referring to Table 1, it is clear that in the MAB a number of correlations were quite modest although the overall battery evidenced a strong *g*, as well as significant verbal and performance lower-order factors.

The implications of the need for task variety are devastating for the available animal literature. Studying a small number of tasks taken from only one or two information modalities as the typical animal study has done may tell us *nothing* about the presence or absence of a factor structure, and, consequently, *g*. Only a carefully-constructed matrix of tasks can reveal that structure. For the studies in Table 2 there is no study that sufficiently marked any factor other than perhaps a spatial factor, let alone included a sufficient variety of tasks to observe the correlation between factor scores—the necessary requirement for marking *g* within a hierarchical structure

Task Complexity

A standard definition of *g* is that it is a "recurrent statistical quantity arising from the fact that every test of any *complexity* contains common elements" (Thompson, 1954, p. 29; emphasis added). This sort of definition reflects the finding in human intelligence that individual differences in *g* covary more reliably with performance on complex tasks than on simple tasks. That is, the *g*-loading of tasks increases with task complexity (Jensen, 1985; similar observations have been made in related situations; see Estes, 1970, for example).

It is exceedingly difficult to apply this criterion to the available animal literature. Given the lack of interest historically in individual differences and in battery construction there are no standard metrics against which the complexity of a task for an animal can be assessed. In human intelligence testing this issue has long been of importance, and as a consequence every standard intelligence test has the age-normed results from large standardization samples against which to determine the relative difficulty level of individual subtests. Nothing similar is available in the animal literature. There is, however, the recognition that previous work on animal *g* may have relied on comparatively simple tasks. In the quote by Warren (1961) that begins this chapter, his conclusion about the failure to observe *g* in animals referred specifically to the fact that "there is no evidence of general intelligence within species [on] *simple association tasks*" (p. 42; parenthetical and emphasis added).

The failure to use tasks that in contemporary work would be regarded as complex is largely a matter of historical timing. When the early work was conducted on this problem in the 1930s many investigators interested in animal learning worked under the assumption that an understanding of more complex functioning could be derived from knowledge of the putatively simple processes of Pavlovian and operant conditioning (see Skinner, 1950, for a rationale for this approach). While this approach was in many ways fruitful there was a growing dissatisfaction with its limitations in leading to an understanding of more complex processes that coincided with the overall rise in cognitive psychology in the late 1960s. The combination of these factors led to a surge of interest in assaying complex cognitive processes in animals. The first volume summarizing these trends appeared in 1978 (viz., Hulse, Fowler, & Honig). Thus, tasks that are now standardfare in the study of animal cognition were not available as research tools when most of the work on animal *g* was done and the conclusion drawn that an animal *g* did not exist.

There is only one study in Table 2 that deliberately used complex tasks and studies problem solving across two modalities. Rajalakshmi and Jeeves (1968) reported significant correlations across a complex spatial task (a Hebb-Williams maze which requires subjects to navigate 12 different maze problems) and a visual task (a pattern discrimination and reversal in rats). Interestingly, they noted that earlier work using simpler visual problems had found zero-order correlations. Livsey (1970) also reported positive correlations between complex tasks including Hebb-Williams and double alternation problems, but used a quite small sample size across three different species.

WHAT IS KNOWN ABOUT ANIMAL *g*?

While the conclusion has regularly appeared in the literature that general intelligence does not exist in animals, an examination of the available literature in light of the requirements needed to assay animal *g* reveals a rather limited and technically-flawed empirical basis for this assertion (see Mackintosh, Wilson, & Boakes, 1985; Macphail, 1982; McGuire & Hirsh, 1977; Scott & Fuller, 1965; Wahlston, 1978, for assertions that animal *g* does not exist). This conclusion does not rule out the possibility that the available literature, however flawed, may reveal some hints as to the nature of animal intelligence. Perhaps the most salient finding emerging from this literature is that most studies have regularly reported substantial correlations between maze tasks. At the least, therefore, animal intelligence can be said to include some sort of spatial factor, although it is not clear whether this factor is general across a variety of spatial procedures. There is some suggestion to this effect in some studies (e.g., Thorndike, 1935), but the literature is too incomplete even with respect to spatial tasks to decide whether the reported correlations are marking

a high-level organizing factor that serves perhaps as the nonhuman analog of the second-order performance factor identified in human intelligence. Interestingly, given that inbred animals were most often used in the early work that demonstrated the presence of this factor, it would appear that this factor may be organized experientially, not genetically (see Anastasi, 1986, for the idea of experiential organization of human abilities).

Given the presence of this spatial factor, it might be asked whether evolution would have favored the organization of behaviors in the form of one factor without the development of others. Does it make sense, in other words, that spatial learning and memory would be integrated to some degree, evidencing some type of spatial *g* so to speak, but that nonspatial problem-solving would consist only of mechanisms specific to particular tasks? If one form of learning has been so organized, is it not reasonable that others have been? And, to extrapolate the argument a speculative step further, would it then not be reasonable to suppose that evolution would have favored a more general organizing factor beyond these lower-order factors, an animal *g*?

Insofar as "factors" are defined as emerging from the study of systematic variance across tasks, the traditional animal learning/cognition literature is of course entirely silent about this possibility (see Eysenck, 1987, and Humphreys, 1987, for similar comments). That same silence applies as well to the contemporary literature. Indeed, Tryon's earlier-cited statement about the lack of interest in individual differences in animals might well be made today as well. A recent volume entitled *Animal Intelligence* (Weiskrantz, 1985), for example, did not devote a single paper to individual differences in animal abilities. It is not clear, as a result, that the term "intelligence" as used with respect to animal behavior bears any relationship to the same term as used to represent the structure of human abilities. The term "animal intelligence," though widely used, has become synonymous and interchangeable with the term "animal cognition," that is, with the investigation of specific tasks such as serial position learning (Swartz, Chen, & Terrace, 1991) or visual concept formation (Herrnstein, 1984).

This difference between the human and animal intelligence literatures represents a remarkable divergence in interests. It might reasonably be thought that the two literatures might articulate with each other, and, indeed, much of the work on animal cognition over the past two decades has underscored the similarities in function between humans and animals (e.g., Hulse, 1993; Wasserman, 1993). Yet, nearly a century after Spearman first identified *g* in humans, there has not been a single study that meets the technical requirements to determine whether *g* exists in animals. The failure to pursue this line of work with animals leaves the animal literature on intelligence sadly lacking a framework for understanding the structure of animal intelligence. Factor analytic work in human intelligence can provide a format for understanding

the connections between various cognitive tasks, how, for example, different kinds of memory are related, or whether all verbal tasks are essentially tapping the same underlying ability. It might in that sense also provide a basis for choosing certain experimental problems for study, thereby linking what Cronbach once termed psychology's two disciplines, the experimental (between-groups) and the correlational (individual differences). Nothing of the sort can be said in the study of animal intelligence. Perhaps that is one reason why Jerison (1985) recently noted that "there is no consensus on the nature of animal intelligence despite a century of research" (p. 21).

The point should be made that these limitations in the available animal literature cannot be taken to suggest that "there must be something there." Absence of evidence is not evidence of absence, but it is also not an indication of a smoking gun. This caution aside, the possibility that a suitably designed research program might reveal an animal *g* would appear to be of significant value to those interested in comparative intelligence. It would at once provide new insights concerning continuities between animals and humans, and would as well provide an animal model of general intelligence that would allow the sorts of experimentation that are impossible in humans.

If such a research program were to reinforce the prevailing view that an animal *g* does not exist, the implications would be equally important. The age-old speculations about distinctions between human and animal minds have in contemporary discourse given way to more precise arguments that center on the presence or absence of language and other forms of symbolic behavior and communication in nonhumans (cf. Chomsky, 1975; Jolly, 1991). That difference should be recognized as in large measure part of the argument concerning human and animal *g*. Indeed, the two sets of distinctions may be linked: the presence of *g* in humans may be, as the result of common linguistic mediation of many human abilities; the absence of *g* in animals may be the result of the lack of a comparable mechanism that functions to allow common mediation of a wide variety of learning situations. Naturally, if an animal *g* were to be found, it would suggest, though not demonstrate conclusively, that the *g* found in humans may be of non-linguistic origins, and common across species rather than unique to ourselves.

NOTES

1. The singular IQ score represents an average of all subtest scores, and, as an average, it is the most highly *g*-loaded measure obtainable from an intelligence battery. For present purposes IQ is almost interchangeable with the notion of *g* as used in this chapter, although it should be recognized that total IQ scores contain sources of variance unique to factors other than *g*.

2. Cattell, in a large series of investigations (e.g., 1987), has argued that *g* itself can be divided into two components, fluid and crystallized intelligence (*gf* and *gc*, respectively). Fluid intelligence is close in concept to Spearman's *g* in that it is presumed to be biologically-based and represents

a general problem-solving skill unconnected to an individual's particular cultural context, whereas *gc* refers to culturally-dependent skills, a kind of "realized" *g* that may take different forms in different cultures. Other workers have identified *gf* and *gc* as much akin to the verbal and performance second-order factors in the model presented in Figure 1 (e.g., Jensen, 1989).

3. Heritability, a forever controversial topic, refers to the extent to which individual differences in test scores are related to genetic differences between individuals (see Locurto, 1991, for a discussion). While the notion of heritability usually refers to total IQ scores, heritabilities can also be determined for individual subtests within a battery, or for that matter, for specific cognitive tasks apart from IQ batteries (e.g., Plomin, 1988). The point made here is that different tasks or subtests exhibit different heritabilities, and these different heritabilities are correlated with a task's *g*-loading.

4. These correlations are actually negative in sign, meaning that faster reaction times are associated with higher IQ scores.

5. The sample is representative not exhaustive because it is difficult to know whether all possible studies have been identified. Standard literature searches reveal no studies of animals within the last 20 years with keywords that label them as studying general intelligence, although a given study may contain correlations that are relevant to this problem. The studies in the table are those that are most often cited in drawing conclusions about *g*, even though many of them were not directed at this problem. An exception is the study by Thompson, Crinella and Yu (1987) who claimed to have documented an animal *g*. Their test battery consisted of a detour problem, a visual discrimination, and two other spatial-type problems. The unusual feature of their work was that different subgroups of rats underwent bilateral lesions to selected brain areas. The sham-operated controls whose data are presented in Table 2 evidenced zero-order correlations across tasks, whereas the operated animals evidenced significant correlations and a robust first factor. The brain-lesioned animals evidenced much greater variances than did controls, suggesting that the lesions had specific behavioral effects, elevating or depressing performance on particular tasks. In a sense, a positive manifold was experimentally-induced in these animals. These data, then, cannot be used to validate the existence of a general factor in intact animals (see also, Crinella & Yu, 1995).

6. Some investigators have used comparisons between selected inbred stains as an alternative to studying differences between individuals (e.g., Wahlsten, 1972). This work may not be strictly relevant to the study of animal *g* because the assumption is that the chosen strains differ reliably on the genetic underpinnings of *g*. It is also problematic to use inbred strain comparisons to mark *g* because inbreeding per se may lower general learning ability (Wahlsten, 1978), and inbred strains often bring sensory and/or motivational anomalies to the learning situation as a byproduct of inbreeding.

7. It might be argued that learning tasks of the sort used in animal work are different from human intelligence tests in that the latter that do not usually require the acquisition of new behaviors but rather assess the application of previous learning to new situations. Jensen (1989) presents evidence that the *g* obtained from intelligence tests and a general factor obtained from learning tasks are virtually identical in humans.

REFERENCES

Anastasi, A., Fuller, J. L., Scott, J. P., & Schmitt, J. R. (1955). A factor analysis of the performance of dogs on certain learning tests. *Zoologica*, *40*, 33-46.

Anastasi, A. (1986). Experiential structuring of psychological traits. *Developmental Review*, *6*, 181-202.

Anderson, B. (1993). Evidence from the rat for a general factor that underlies cognitive performance and that relates to brain size: Intelligence? *Neuroscience Letters, 153*, 98-112.

Bitterman, M. E. (1965). The evolution of intelligence. *Scientific American*, *212*, 92-100.
Bitterman, M. E. (1969). Thorndike and the problem of animal intelligence, *American Psychologist*, *24*, 444-453.
Bitterman, M. E. (1975). The comparative analysis of behavior. *Science*, *188*, 699-709.
Blaha, J., & Walbrown, F. H. (1991). Hierarchical factor structure of the Wechsler Preschool and Primary Scale of Intelligence-Revised. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *3*, 455-463.
Bouchard, T. J. Jr. (1993). The genetic architecture of human intelligence. In P. E. Vernon (Ed.), *Biological approaches to the study of human intelligence* (pp. 33-93). Norwood, NJ: Ablex.
Brody, N. (1992). *Intelligence* [*2nd. Ed.*]. New York: Academic Press.
Burt, C. (1949). Alternative methods of factor analysis and their relations to Pearson's method of 'principal axes.' *British Journal of Psychological Statistics, Section 2*, 98-121.
Campbell, A. A. (1935). Community of function in performance of rats on alley mazes and Maier reasoning apparatus. *Journal of Comparative and Physiological Psychology*, *31*, 225-235.
Cardon, L. R., Fulker, D. W., DeFries, J. C., & Plomin, R. (1992). Multivariate genetic analysis of specific cognitive abilities in the Colorado Adoption Project at age 7. *Intelligence*, *16*, 383-400.
Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. New York: North-Holland.
Child, D. (1990). *The essentials of factor analysis* [*2nd .Ed.*]. London: Cassell Educational Ltd.
Chomsky, N. (1975). *Reflections on language*. New York: Pantheon.
Creel, D. (1980). Inappropriate use of albino animals in research. *Pharmacology, Biochemistry & Behavior*, *12*, 969-972.
Commins, W. D., McNemar, Q., & Stone, C. Intercorrelations of measures of ability in the rat. *Journal of Comparative and Physiological Psychology*, *14*, 225-235.
Crinella, F. M., & Yu, J. (1995). Brain mechanisms in problem solving and intelligence: A replication and extension. *Intelligene, 21*, 225-246.
Dunlap, J. W. (1933). The organization of learning and other traits in chickens. *Comparative Psychological Monographs*, *8*(4).
Estes, W. (1970). *Learning theory and mental development*. New York: Academic Press.
Eysenck, H. J. (1979). *The structure and measurement of intelligence*. New York: Springer-Verlag.
Eysenck, H. J. (1987). The several meanings of intelligence. *Brain and Behavioral Sciences*, *10*, 663.
Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
Gottfredson, L. S. (1986). Societal consequences of the *g* factor in employment. *Journal of Vocational Behaviour*, *29*, 379-410.
Gould, S. J. (1981). *The mismeasure of man*. New York: W. W. Norton.
Gustafsson, J. E. (1988). Hierarchical models of individual differences. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 4). Hillsdale, NJ: Erlbaum.
Hebb, D. O., & Williams, K. (1946). A method of rating animal intelligence. *Journal of Genetic Psychology, 34*, 59-65.
Herrnstein, R. J. (1984). Objects, categories, and discriminative stimuli. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition*. Hillsdale, NJ: Erlbaum.
Hulse, S. H. (Ed.). (1993). Symposium on animal cognition, *Psychological Science*, *4*, 154-189.
Hulse, S. H., Fowler, H., & Honig, W. K. (Eds.) (1978). *Cognitive processes in animal behavior*. Hillsdale: Erlbaum.
Humphreys, L. G. (1987). Psychometric considerations in the evaluation of intraspecies differences in intelligence. *Brain and Behavioral Sciences*, *10*, 668-669.
Jensen, A. R. (1981). *Straight talk about mental tests*. New York: The Free Press.
Jensen, A. R. (1985). The nature of black-white difference on various psychometric tests: Spearman's hypothesis. *The Behavioral and Brain Sciences*, *8*, 193-263.
Jensen, A. R. (1986). *g*: artifact or reality? *Journal of Vocational Behavior*, *29*, 301-331.

Jensen, A. R. (1989). The relationship between learning and intelligence. *Learning and Individual Differences*, *1*, 37-62.

Jensen, (1992). Understanding *g* in terms of information processing. *Educational Psychology Review, 4,* 271-308.

Jensen, A. R., & Weng, L. (1994). What is a good *g*? *Intelligence*, *18*, 231-258.

Jerison, H. J. (1985). Animal intelligence as encephalization. In L. Weiskrantz (Ed.), *Animal intelligence* (pp. 21-36). Oxford: Clarendon Press.

Jolly, A. (1991). Conscious chimpanzees? A review of recent literature. In C. A. Ristau (Ed.), *Cognitive ethology: The minds of other animals* (pp. 231-252). Hillsdale, NJ: Erlbaum.

Kamin, L. J. (1973). *The science and politics of IQ*. Potomac, MD: Erlbaum.

Kline, P. (1991). *Intelligence: The psychometric view*. London: Routledge.

Liggett, J. R. (1925). A note on the reliability of the chick's performance in two simple mazes. *Journal of Genetic Psychology*, *32*, 470-480.

Livsey, P. J. (1970). A consideration of the neural basis of intelligent behavior: Comparative studies, *Behavioral Science*, *15*, 164-170.

Locurto, C. (1991). *Sense and nonsense about IQ: The case for uniqueness*. New York: Praeger.

Mackintosh, N. J. (1988). Approaches to the study of animal intelligence. *British Journal of Psychology*, *79*, 509-525.

Mackintosh, N. J., Wilson, B., & Boakes, R. A. (1985). Differences in mechanisms of intelligence among vertebrates. In L. Weiskrantz (Ed.), *Animal intelligence* (pp. 53-66). Oxford: Clarendon Press.

Macphail, E. M. (1982). *Brain and intelligence in vertebrates*. Oxford: Clarendon Press.

Macphail, E. M. (1985a). Vertebrate intelligence: the null hypothesis. In L. Weiskrantz (Ed.), *Animal intelligence* (pp. 37-50). Oxford: Clarendon Press.

Macphail, E. M. (1985b). Comparative studies of animal intelligence: Is Spearman's *g* really Hull's D? *Behavioral and Brain Sciences*, *8*, 234-235.

Macphail, E. M. (1987). The comparative psychology of intelligence. *Behavioral and Brain Sciences*, *10*, 645-695.

McGuire, T. R., & Hirsch, J. (1977). General intelligence (*g*) and heritability (H^2, h^2). In I. C. Uzgiris & F. Weizmann (Eds.), *The structuring of experience* (pp. 25-72). New York: Plenum.

Miller, L. T., & Vernon, P. A. (1992). The general factor in short-term memory, intelligence, and reaction time. *Intelligence*, *16*, 5-30.

Nettleback, T. (1987). Inspection time and intelligence. In P. A. Vernon, (Ed.), *Speed of information-processing and intelligence* (pp. 295-346). Norwood, NJ: Ablex.

Plomin, R. (1988). The nature and nurture of cognitive abilities. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 4, pp. 1-33). Hillsdale, NJ: Erlbaum.

Plomin, R., DeFries, J. C., & McClearn, G. E. (1990). *Behavioral genetics: A primer* (2nd. ed.). New York: W. H. Freeman.

Rajalakshmi, R., & Jeeves, M. A. (1968). Performance on the Hebb-Williams Maze as related to discrimination and reversal learning in rats. *Animal Behavior*, *16*, 114-116.

Ree, M. J., & Earles, J. A. (1991). The stability of *g* across different methods. *Intelligence*, *15*, 271-278.

Royce, J. R. (1950). The factorial analysis of animal behavior. *Psychological Bulletin*, *47*, 235-259.

Royce, J. R. (1966). Concepts generated in comparative and physiological psychological observations. In R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology* (pp. 642-683). Chicago: Rand Mcnally.

Scott, J. P., & Fuller, J. L. (1965). *Dog behavior: The genetic basis*. Chicago: University of Chicago Press.

Seligman, M. (1970). On the generality of the laws of learning, *Psychological Review*, *77*, 406-418.

Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review*, *57*, 193-216.

Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York: Macmillan.
Sternberg, R. J. (1986). *The triarchic mind: A new theory of human intelligence*. New York: Viking.
Sternberg, R. J., & Gardner, M. K. (1982). A componential interpretation of the general factor in human intelligence. In H. J. Eysenck (Ed.), *A model for intelligence*. New York: Springer.
Swartz, K. B., Chen, S., & Terrace, H. S. (1991). Serial learning by Rhesus monkeys: I. Acquisition and relation of multiple four- item lists. *Journal of Experimental Psychology: Animal Behavior Processes*, *17*, 396-410.
Taylor, R. L., & Ziegler, E. W. (1987). Comparison of the first principal factor on the WISC-R across ethnic groups. *Educational and psychological measurement*, *47*, 691-694.
Terman, L. M. (1916). *The measurement of intelligence*. Boston: Houghton Mifflin.
Terman, L. M., & Merrill, M. A. (1937). *Measuring intelligence: A guide to the administration of the new revised Stanford-Binet tests of intelligence*. Boston: Houghton Mifflin.
Thompson, G. H. (1954). *The geometry of mental measurement*. London: University of London Press.
Thompson, R., Crinella, F. M., & Yu, J. (1987). *Brain mechanisms in problem solving and intelligence: A lesion survey of the rat brain*. New York: Plenum Press.
Thorndike, E. L. (1911). *Animal intelligence: Experimental studies*. New York: Macmillan.
Thorndike, E. L. (1935). Organization of behavior in the albino rat. *Genetic Psychology Monographs, 17*(1).
Thorndike, R. L. (1987). Stability of factor loadings. *Personality and Individual Differences*, *8*, 585-586.
Thorndike, R. L. (1994). *g*. *Intelligence*, *19*, 145-155.
Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.
Tomlin, M. I., & Stone, C. P. (1934). Intercorrelations of measures of learning ability in the albino rat. *Journal of Comparative and Physiological Psychology*, *17*, 73-88.
Tryon, R. C. (1942). Individual differences. In F.A. Moss (Ed.), *Comparative psychology* (pp. 330-365). Englewood Cliffs: Prentice-Hall.
Vernon, P. E. (1965). *The structure of human abilities*. London: Methuen.
Vernon, P. A. (1987). *Speed of information-processing and intelligence*. Norwood, NJ: Ablex.
Wahlsten, D. (1972). Genetic experiments with animals: A critical review. *Behavioral Biology*, *7*, 143-182.
Wahlsten, D. (1978). Behavioral genetics and animal learning. In H. Anisman & G. Bignami (Eds.), *Psychopharmacology of aversively motivated behavior* (pp. 63-118). New York: Plenum.
Warren, J. M. (1961). Individual differences in discrimination learning by cats. *Journal of Genetic Psychology*, *98*, 89-93.
Warren, J. M. (1977). A phylogenetic approach to learning and intelligence. In A. Oliverio (Ed.), *Genetics, environment and intelligence* (pp. 37-56). New York: North-Holland.
Wasserman, E. A. (1993). Comparative cognition: Beginning the second century of the study of animal intelligence. *Psychological Bulletin*, *113*, 211-228.
Weiskrantz, L. (Ed.). (1985). *Animal intelligence*. Oxford: Clarendon Press.

INFANT COGNITION AND GENERAL INTELLIGENCE

D. Wayne Mitchell and John Colombo

Some psychologists may be born to study intelligence; others have intelligence thrust upon them. The authors of this chapter fall clearly into the latter of these two categories. Our own research interests have focused primarily on individual differences in cognitive processes in early infancy, and how these processes may (or may not) contribute to the quality of cognitive function observed later in life.[1] On occasion, this research necessitates comparing the relationship of infants' performance on perceptual-cognitive tasks with subsequent cognitive function in childhood. As a result, psychologists in this line of inquiry are called upon to administer, interpret, and ponder the significance of standardized intelligence tests. In turn, many developmentalists like us have found themselves in the position of being asked to comment about research and theory on the topic of intelligence. We find ourselves in just such a position here.

Indeed, a serious theoretical consideration of intelligence in infancy causes one to contemplate the fundamental assumptions underlying the construct of intelligence. This is particularly the case for the concept of "general intelligence"

Advances in Cognition and Educational Practice, Volume 4, pages 101-118.

ISBN: 0-7623-0105-8

(*g*), which seems to fuel the core of many of the academic disputes within the field. Although *g* has not been widely discussed within the context of early cognitive and intellectual development, the construct has influenced work in the area of early standardized testing and infant cognition. It is our aim in this chapter to trace that path of influence, and to consider the ultimate advantages and disadvantages of this influence. We begin with a brief historical account of early infant intelligence tests and their association with the concept of *g*, and how findings from these early tests were brought to bear on that concept. We then recount recent work on the continuity of intelligence from infancy, and how that work has brought *g* to the brink of consideration in the developmental literatures. We conclude with a discussion of what such a consideration might contribute to our understanding of infant intelligence, and conversely, to what degree the study of cognitive and intellectual function in infancy may contribute to our understanding of *g*.

INTELLIGENCE IN ADULTS AND INFANTS: A BRIEF HISTORICAL SKETCH

Owing to the work of Galton and Binet in Europe, the topic of adult intelligence had become a distinct focus of inquiry within psychology in the United States by the turn of the twentieth century (see, e.g., Cattell, 1890). During the period coinciding with World War I, psychometricians had undertaken large-scale efforts toward the development of standardized measures of adult intellectual performance.

Part of this effort was motivated by a practical desire to identify the level of intellectual ability in particular individuals (e.g., various tests were developed for the assessment of military personnel; see, e.g., Yerkes, 1921). Another distinct aim of some of this work, however, was to identify and understand the basic structure of intelligence. Typically, this question was addressed via factor analyses of subtests or subdomains of the intelligence tests themselves. Of particular interest was the question of whether intelligence test performance reflected a set of distinct and independent skills, or whether such performance could be reduced to a single general factor. The fundamental dichotomy of this debate can be traced to the theoretical differences that characterized Galton (1892) and Binet (1903). While Galton focused on the biological underpinnings and evolutionary history of exceptional mental traits, Binet's emphasis concerned the adaptation of specific skills to particular environments as a method of accounting for the performance of individuals spanning a wide range of abilities.

The tradition of Galton was extended by the hypothesized existence of a biologically-based general factor by Spearman (1904), while Binet's emphasis was subsequently reflected in Thurstone's (1938) multiplicative generation of over one hundred "primary mental abilities" (see also Guilford, 1967).

Eventually, participants on both sides of the debate acknowledged the need for both general and specific factors in a unified model of human intellectual function (Spearman & Jones, 1950; Thurstone & Thurstone, 1941). However, the question of whether intellectual function was best characterized as a general or modular phenomenon survives in varied incarnations (see, e.g., Gardner, 1983; Sternberg, 1985a), and is still the source of lively discussion (cf. Detterman, 1992).

The Normative Infant Tests

After the first World War, the topic of intelligence was applied to developmental issues as well; standardized psychometric instruments for the assessment of infants and young children began to appear shortly thereafter. Indeed, the first wave of infant intelligence scales[2] appeared in the 1920s and 1930s (Bayley, 1933; Fillmore, 1936; Gesell, 1925; Shirley, 1933; Stutsman, 1931). These infant tests were developed largely under the auspices of the various Child Welfare Institutes that were established just after World War I, and their format reflected the normative strategies that were popular within the discipline at the time and which broadly characterized the study of human child development at those institutes (for reviews, see Colombo, 1993; Brooks-Gunn & Weinraub, 1983).

In keeping with the normative tradition for studying individual differences, the structure of these early "traditional" infant tests closely resembled that of the adult tests; the difficulty of items progressively increased in difficulty throughout the test, and performance on higher-order items was assessed both in terms of the ability to complete the task and the rapidity with which it was completed. Most importantly, on both the adult and infant assessments, the highest items passed were taken as the measure of the individual's intellectual "level," which was scaled with reference to normative values obtained from the population as a whole.

INFANT "INTELLIGENCE" TESTS: A SEARCH FOR CONSTANCY?

Judging by the proliferation of the scientific literature on this topic during the 1920s and 1930s, the industry surrounding the development and standardization of these new assessments was quite vigorous. What motivated the construction and use of the normative infant intelligence tests? Although a definitive answer will not be forthcoming here, it does appear that the perceived practical success achieved in adult intelligence testing, coupled with the simultaneous appearance of the Binet tests for children (Binet & Simon, 1905, 1908, 1911; see also Binet & Henri, 1895) in the United States (e.g., Terman,

1916), may well have contributed to the belief that "intelligence" might be successfully assessed in very young children and infants.[3]

This belief, however, had to have rested on the logical assumption that the construct assessed by intelligence tests was, in the terminology of the early differential psychologists (Nemzek, 1933; Thorndike, 1940), "*constant*" across repeated testings administered at different periods during the life span (see Anastasi, 1969). Of relevance to the present discussion is the fact that the presence of such constancy (*especially* from infancy onward) would constitute relatively strong support for the existence of *g*. Thus, we contend that the conceptualization of intelligence as a general and unchanging factor did, in fact, contribute to the rationale for, and construction of, these initial assessments of "mental" or "intellectual" ability for infants and children.

INFANT AND ADULT INTELLIGENCE TESTS: EMPIRICAL RELATIONSHIPS

The initial empirical evaluations of such constancy in intellect across the lifespan, however, were not particularly encouraging. Indeed, data from the earliest longitudinal studies that examined the predictive validity of these infant tests (Furfey & Muehlenbein, 1932; Stott & Ball, 1965; Thorndike, 1940) suggested that performance on infant intelligence tests did not correlate with performance on intelligence tests administered later during the lifespan.

This was a replicable finding; indeed, over the subsequent decades, it became widely accepted that infants' performance on variants of the early infant intelligence tests predicted well neither to childhood nor to adult intelligence test performance. The most comprehensive reviews on the predictive validity of these tests came from McCall (1979b, 1983). He reported correlations of +.12 and +.26 between infant tests administered respectively at 0 to 6 months of age and 7 to 12 months of age with childhood intelligence test performance spanning ages from 8 to 18 years. Although the infant tests proved to be relatively useful for diagnostic and clinical practices with high-risk samples of infants (e.g., Honzik, 1976), they generally did not predict to later outcome for infants who fell within the normal or above-normal range of developmental status.

Non-prediction and Intelligence: Implications

Several explanations were posited for the failure of the infant tests to predict to childhood and adult tests of intellectual performance. Some were purely statistical or psychometric in nature; Cattell (1987) argued that research on the infant tests lacked sufficient statistical power for factor analytic studies of structural content because their standardization sample sizes were too small. Additionally, Cattell (1987) argued that the items chosen to be included in the

early standardized infant tests were not constructed to map directly onto the abilities presumed to be assessed on the adult tests; indeed, the item content and domains of the infant and childhood/adult tests are markedly different.

Other theorists, however, focused on a more abstract possibility. For these scientists, the failure of prediction was attributed to a violation of the basic assumption of constancy of intellectual function from infancy (McCall, Hogarty, & Hurlburt, 1972; Stott & Ball, 1965). The independence of performance on the infant and child/adult intelligence tests was interpreted as an indication that the developmental course for intelligence was somehow "discontinuous" (e.g., McCall, 1979a); most importantly for the present discussion, the lack of prediction was difficult to reconcile with a presumably robust general factor of intelligence. For example, Stott and Ball (1965) wrote:

> The fact that the results of infant tests bear no consistent relation with later test results is, of course, especially "damaging" to the idea that mental development is purely a matter of quantitative growth (p. 134).

About a decade later, in positing a five-factor model of infant intelligence, McCall et al. (1972) stated that:

> The overriding implication . . . is that a simple conception of a constant and pervasive *g* factor is probably not tenable as a model for "mental" development, especially for the infancy period . . . to label as "mental" performances at every age perpetuates the belief in a pervasive and developmentally constant intelligence . . . The network of transitions between skills at one age and another is likely more specific and complex than once thought, and not accurately subsumed under one general concept (p. 746).

"CONTINUITY" IN MENTAL DEVELOPMENT

During the 1980s, claims of discontinuity and instability in intelligence from infancy were tempered by findings from basic research on infant cognitive function. During the 1960s and 1970s, various methods were developed for assessing fundamental cognitive abilities in preverbal infants. These abilities included variables thought to reflect stimulus encoding (i.e., from the visual habituation paradigm; see Horowitz, Paden, Bhana, & Self, 1972), and visual discrimination/recognition memory (from the paired-comparison paradigm; Fagan, 1971, 1974). During the late 1970s and 1980s, several of these infant cognition measures were incorporated into longitudinal designs (e.g., Miller et al., 1980). The results of these longitudinal studies were unexpectedly and remarkably consistent: individual differences in these measures of infant cognition predicted modestly, but significantly, to measurements of intelligence later in childhood.

This literature, first reviewed by Bornstein and Sigman (1986), has been reassessed a number of times since (e.g., Colombo, 1993; Colombo & Mitchell, 1990; McCall, 1994; McCall & Carriger, 1992; Rose, 1989), with the same general conclusion: individual differences in measures of stimulus discrimination and encoding from infancy are observed to account for between 12 percent and 25 percent of the variance (i.e., correlations from .35 to .50) in later performance on intelligence or cognitive tests in childhood. Later reviews (e.g., Colombo, 1993) suggested that individual differences in infants' long-term retention (Fagen & Ohr, 1990) and oculocephalic reaction time (e.g., DiLalla et al., 1990) also significantly predicted to later intelligence; the magnitude of these correlations roughly approximate the level of those for stimulus encoding and recognition memory.

Thus, measures of perceptual-cognitive processes taken during the first six months of life predicted modestly but significantly to intellectual and cognitive status in later childhood. Indeed, the longest running longitudinal studies to date indicate prediction from early infancy to preadolescence (Sigman, Cohen, Beckwith, Asarnow, & Parmelee, 1991; Rose & Feldman, 1995). Although the infant cognition measures account for a limited amount of the variance in childhood intelligence scores, the amount accounted for by these measures represents a ten- to twenty-fold increase over that accounted for by the early infant intelligence tests. Furthermore, it is likely that the level of prediction from infancy measures to childhood IQ is constrained by the low reliabilities of the infant measures; thus, the level of prediction may in fact be an underestimate of the variance that such measures may ultimately account for in later IQ (Colombo, 1993). These facts have been taken to suggest that these measures may tap some very basic emergent characteristics of the cognitive/intellectual system.

Furthermore, whatever these aspects may be, they appear to remain relatively constant, or "continuous"[4] across a part of the life span that is characterized by rapid change, tumultuous spurts in growth, and numerous behavioral "reorganizations." As a result of these findings, there has been some interest in the development and standardization of instruments using these measures for use in early classification and diagnostic procedures, particularly for identifying infants at risk for later cognitive deficits (e.g., Fagan, Singer, Montie, & Shepard, 1986). As a result of this interest, the Fagan Test of Infant Intelligence (FTII; Fagan & Shepard, 1986/1987) has been developed and made commercially available.

Some Current Issues in the Continuity Literature

We would be remiss here simply to report a summary of the prediction-from-infancy phenomenon without reference to some important controversies that surround it. The first involves the aforementioned unreliability of the infant

measurements. Indeed, the infant cognition measures show predictive validity despite the fact that they do not possess the kind of test-retest reliability or internal consistency that would satisfy most psychometricians whose frame of reference is the well-standardized and highly reliable adult tests (see Benasich & Bejar, 1992; Fagan & Detterman, 1992). This lack of reliability constitutes a real impediment to the actual predictive power of the measures (McCall & Carriger, 1992), and further, it tends to inspire less confidence in potential academic or clinical consumers with respect to the robustness and precision of what is being measured (Benasich & Bejar, 1992; McCall, 1994).

A second issue involves the question of whether prediction from these measures is linear or nonlinear. Each of the measures of infant cognition discussed previously can be scored on a continuous scale, but the data to this point are equivocal as to whether variance along the entire range of the continuum does in fact carry predictive value. Rose, Feldman, Wallace, and McCarton (1989) and McCall and Carriger (1992) have put forth a case for linear prediction; that is, continuous variation along the scale of, for example, novelty preference does in fact add to the predictive power of the novelty preference measure. However, Fagan et al. (1986) have demonstrated the predictive validity of novelty preference for various samples of infants by dichotomizing the groups into "non-risk/risk," or "normal/suspect" classifications. Some interpretive guidelines for the FTII (Fagan & Shepard, 1986/1987) also suggest the utility of a dichotomous criterion. Similarly, we have tended toward an interpretation of infant fixation duration (one of the measures that is predictive of later intelligence; see Rose, Slater, & Perry, 1986; Sigman et al., 1991) that is closer to this latter interpretation. This issue will likely become more important if these measures are to become more widely used for diagnostic and/or clinical purposes.

A third issue of contention concerns whether this prediction from infancy is limited to measures that involve visual processing, or whether the functions being tapped are more centrally-mediated. To this point, a vast majority of those studies documenting prediction of childhood intelligence from infancy have employed measures that involve visual processing (see, however, O'Connor, Cohen, & Parmelee, 1984). It is thus appropriate to ask whether the prediction phenomenon is limited to, or constrained by, some aspect of visual function, or whether it involves a more general, centrally-mediated mechanism (Colombo, 1993). More general, central mechanisms, such as the response to novelty (Berg & Sternberg, 1985) or inhibition (McCall, 1994; McCall & Carriger, 1992), have been posited to account for the mechanisms underlying prediction from infancy, but some specific models have been put forth that fundamentally involve only the pathways for geniculostriate vision, or those that mediate visuospatial attention and search (Colombo, 1995).

Prediction and Intelligence: Implications

Whereas the lack of prediction of childhood and adult intellectual status from infancy with traditional normative tests was taken as a lack of support for the concept of *g*, the recent findings of significant prediction from infancy with laboratory-based measures of infant cognition may well be taken as empirical support for a hypothesized general factor of intelligence (see also Fagan, 1984a). Indeed, some of the extant evidence on the mechanisms underlying prediction suggests that the infant measures converge on an underlying construct, rather than simply reflecting constancy on a fundamental and isolated cognitive ability.

For example, infant novelty preferences (which comprise the FTII) are, in laboratory studies, usually taken as a measure of the infant's recognition memory for visual stimuli. Fagan (1984b), however, demonstrated that infant novelty preferences correlated with intelligence test scores later in childhood, but *not* with recognition memory performance later in childhood. This suggested that infant novelty preferences were more strongly related to some underlying construct that was also tapped by intelligence test performance, rather than simply reflecting continuity in recognition memory.

Similarly, several reports from Rose's research program (Rose, Feldman, Wallace, & Cohen, 1991; Rose, Feldman, & Wallace, 1992) have suggested that the predictive variance carried by infant novelty preferences in accounting for later individual differences in intelligence are mediated largely by language ability. In this case as well, the predictive validity of novelty preferences appears to tap a system that is manifest in terms of other cognitive or intellectual products. Indeed, Rose's (e.g., Rose et al., 1991) findings strongly suggest that novelty preference taps some construct that serves as a precursor to verbal ability; this apparently can be tapped during the first 9 months of life, a time when verbal ability itself is not in any way manifest in any overt behavioral skill (see also Rose & Feldman, 1995).

Colombo (1993) reviewed the entire prediction literature in an attempt to deduce the underlying factors responsible for the continuity, and proposed a two-factor model that included speed of processing and memory. This model has received some support in the literature (e.g., Jacobson et al., 1992; Rose & Feldman, 1995).

Thus, the prediction of childhood and adult intelligence from infancy occurs across a relatively unstable portion of the lifespan, one that is punctuated by various behavioral reorganizations (e.g., McCall et al., 1972). Such prediction, furthermore, is observed across what are (at least at face value) widely discrepant and divergent types of cognitive tasks. Finally, the amount of variance accounted for by infant cognition tasks in later intelligence test performance is roughly equal to that commonly observed for the principal component derived from factor analyses of intelligence test performance (i.e.,

the statistical evidence from which the plausibility of *g* is derived). In sum, the findings related to the continuity of intellectual function from infancy are broadly consistent with the existence of some type of general factor within the infant measures, and with the operation of such a general factor from infancy; this point has not been discussed widely in the developmental literatures (see Fagan, 1984a). In the remainder of this chapter, we further explore this possibility, as we consider the relationship between infant intelligence, adult intelligence, and *g*.

INFANT COGNITION AND GENERAL INTELLIGENCE

The concept of g

The concept of *g* has long been a controversial one. This is the case partly because its significance and interpretation depend largely upon one's view of what comprises an intelligence score, and on this point there are many and varied viewpoints.

Keating and MacLean (1987) have noted that the positions on this issue generally fall into two categories. The first holds that an individual's performance on an IQ test reveals something about that individual's level of competence that in some way "explains" individual and developmental differences in performance. Alternately, the second perspective suggests that an IQ test score simply provides information about an individual's current developmental and/or cognitive status, without further reference to a fixed, internal "trait" as an explanatory construct for performance.

It is the first of these two positions that leads logically to an emphasis on the construct of *g*. This position is ultimately attributable to the traditional psychometric notions of intelligence, and its empirical basis is drawn from factor analyses of intelligence test performance. An especially important facet of these analyses is the interpretation of the first principal component and associated factor loadings that typically emerges; the existence and strength of the first principal component has traditionally been thought to reflect the existence and pervasiveness of *g* (see Spearman, 1904; Terman, 1916; see also Jensen, 1979, 1984).

The concept of general intelligence has been contested in two basic ways. The first involves challenges to either the factor analytic studies from which *g* is derived (e.g., Thurstone, 1938, 1947; Guilford, 1967; Horn, 1985) or the validity of the interpretation of those factor analyses in the reification of the hypothetical construct (see also Carroll & Maxwell, 1979; Gould, 1981, 1987). A second approach is to dismiss or call the contribution of the principal component into question, based on the fact that it typically does not account for more than one quarter of the variance in overall individual performance on intelligence tests (see, e.g., Ceci, Nightingale, & Baker, 1992; Sternberg, 1985).

Despite these attacks, however, the concept of *g* in various forms has generally remained viable in the behavioral sciences. In our own analysis, however, the philosophical questions concerning reification and the empirical questions concerning the amount of variance accounted for do not diminish the potential importance of what the principal component represents (see Hunt, 1995). Indeed, a principal component *does* emerge from nearly all the factor analyses of intelligence test performance, irrespective of the investigator's view on the topic. On this component, all of the subtests contribute to an agglomeration that Hunt (1980) has referred to as the "positive manifold." The intercorrelation of performance on the subscales indicates some common variance. Furthermore, because this covariance is consistently observed in such analyses of individual differences in test performance (even by its detractors), we think it constitutes a phenomenon, and is thus worthy of explanation.

Does g Explain the Continuity of Intelligence from Infancy?

It is tempting to suggest that *g* is somehow responsible for the prediction of childhood intelligence from infancy. Indeed, it would be quite easy to close this chapter with the conclusion that the prediction of childhood intelligence from infancy is, in fact, mediated by *g*. In our opinion, however, there is an important drawback to such a course of action.

In the preceding section, we have claimed that the principal component emerging from factor analytic studies of intelligence test performance is worthy of investigation and explanation. But does the application of the label "*g*" to the prinicipal component or "positive manifold" (Hunt, 1980) of intelligence tests really help toward the advancement of understanding the processes that produce it? We would argue that it does not, because the attribution of individual differences in any manifest behavioral performance to a construct label is inadequate; that is, while attribution might suffice as a "shorthand" method of referring to those individual differences, in the absence of the specification of any discrete processes, such attribution simply cannot, and does not, suffice as a real explanation for those individual differences.

We do not reject the notion that some general process (or processes) may contribute to performance on a number of domains that comprise our definitions of intelligence. We would insist, however, that the general factor be necessarily reducible to specific parameters or components of the information processing system. Furthermore, we do not reject the notion that the information-processing parameters or components that give rise to the inference of *g* in the adult may be the same as those parameters or components that can be measured in infancy and which predict to adult intelligence. We would contend, however, that this is not an appropriate stopping place for the science; by simply attributing continuity in intelligence from infancy to *g*, we may never really come to know what these parameters or components are.

Over the past twenty years, the concept of intelligence has been re-examined within the context of information processing models of cognitive performance (e.g., Hunt, 1980, 1995; Pellegrino & Glaser, 1979). These approaches have focused on the individual differences in cognitive function that might underlie the covariance of individuals' performance on subscales of intelligence tests. As such, these recent investigations have placed more emphasis on the "processes" that may underlie performance and be brought to bear as explanations for such performance rather than on the final "product," or "trait" of general intelligence (e.g., Ackerman, 1986, 1988; Ceci et al., 1992; Larson & Alderton, 1992).

If the principal component that gives rise to the postulation of *g* is conceptualized in terms of cognitive processes, then those processes should be fairly basic or fundamental parameters of the information processing system. For example, speed of processing (e.g., Vernon, 1987) and short-term memory capacity (Larson & Alderton, 1992) are often considered as prime candidates for the processes underlying *g*. We have applied this same type of approach to the analysis of infants' performance on attentional and learning tasks (see Colombo, 1993; Colombo & Mitchell, 1990, 1988).

It is of some note that *g* is often considered within contexts where the stability of individual traits is assumed and emphasized (see Eysenck, 1992). With this focus, the construct of *g* has been consistently presented as an end "product" of behavioral function. The attribution of individual differences to such a product shifts the focus away from both the consideration and study of the antecedent conditions that contribute to developmental and individual differences in cognitive processes, which in turn contribute to intelligence test performance (see Snow, 1982; Sternberg, 1985b). We have elsewhere (Colombo & Mitchell, 1990) argued for the advantages in studying "process" over "product" from the point of view of the study of cognition, and toward a more comprehensive understanding of individual differences (see also Underwood, 1975).

In short, we advocate accounting for individual differences in intelligence not by attributing them to a construct, but by investigating the actual, specific, cognitive processes that contribute to them. Our position is not new; for more than a half century, scientists who perceived intelligence more as a byproduct of cognitive function, rather than as an entity in itself, have called for new directions in the study of intelligence (e.g., Gesell, 1925; Resnick, 1976; Sigel, 1963; Tyler, 1976). Indeed, a number of researchers are actively investigating the contribution of different aspects of information processing speed to the general component (e.g., Jensen, 1993; Reed & Jensen, 1992; Vernon & Mori, 1992). We would, however, like to avoid the explicit application of *g* to the infant literatures. While the psychometric testing tradition has its uses and contributions, its theoretical focus is limited from the perspective of building a cognitive science of individual differences.

Can Research on Infant Cognition Contribute to our Understanding of *g*?

In the previous section, we contemplated whether models of adult general intelligence might contribute to our understanding of the continuity of cognitive function from infancy. In this section, we consider the converse question: that is, whether models of infant cognitive/intellectual function might contribute to an understanding of the processes that induce the inference of *g*.

We believe that, in fact, there is good reason for this to be the case. If we consider *g* as simply the covariance that is often observed among subtests of the intelligence test, we would propose that the cognitive processes that contribute that covariance might actually be better accessed or studied in the infant than in the adult. Other authors have noted that when cognitive processes become routine and automatic, there is little to be learned about how these processes are actually executed (Keating, 1984; Sternberg, 1985a). Indeed, Vygotsky (1978) has stated that:

> In psychology we often meet with processes that have already died away, that is, processes that have gone through a very long stage of historical development and have become fossilized...These fossilized forms of behavior are most easily found in the so-called automatic or mechanized psychological processes...They have lost their original appearance and their outer appearance tells us nothing whatsoever about their internal nature (pp. 63-64).

Our point here is that, in the infant, the cognitive processes or parameters that ultimately contribute to intelligence are more likely to be in an emergent or malleable state than they are in the adult. They are in all probability executed more slowly, and involve more cognitive "effort" during infancy than when they are exhibited in adults. Along with the possibility that such functions might be observed in a more emergent state, it is also likely that we will be able to study how they change, and how they are affected by early environmental conditions or experience (e.g., Gottleib, 1976, 1992). We would argue that such characteristics would actually *facilitate* the empirical study of these processes.

In an effort to address this issue, it is likely that there will be a need to develop new methods for research that parallel those employed with adults. In this vein, new paradigms for studying infant cognition have in fact emerged during the past 5 years (e.g., Haith & McCarty, 1990; Johnson, Posner, & Rothbart, 1991). Additionally, there will be a need to more thoroughly understand the already-established paradigms of habituation, paired-comparison, and conditioning, as a thorough understanding of the cognitive tasks in question is a fundamental step toward understanding the processes that contribute toward individual differences in task performance. Furthermore, some effort will need to be expended so that task difficulty for both new and old measures of infant

cognition can be varied with parametric precision; a task that is too easy or too difficult will yield little or no variation in cognitive process (DeLoache, Cassidy, & Brown, 1985; Belmont & Mitchell, 1987).

Finally, this research will require an orientation toward questions of "how" early information processing skills develop during infancy. The "how" question is far from being a new one, and is not entirely independent of the question of "what" develops (e.g., Flavell & Wohlwill, 1969). However, the "how" of development necessarily involves an assessment of the role(s) of experience in the maintenence, facilitation, and induction of cognitive function (Aslin & Dumais, 1980; Gottlieb, 1992).

CONCLUDING REMARKS

In this chapter, we have briefly documented the history of the study of infant intelligence, and some of the implications and controversies present in the current state of the discipline. Additionally, we have considered the evidence for and the role of *g* in both adult and infant intelligence. In conclusion, we emphasize two basic points with regard to what the theory of the structure adult intelligence might have to offer to our understanding of infant intelligence, and what the study of infant intelligence might have to offer to the study of adult intelligence.

The first point is that, although it is quite tempting to attribute the continuity of intellectual function to the abstract construct of *g*, such attribution would explain little and might also inhibit progress toward an understanding of either the continuity of intelligence or *g* itself. A second point is that, although the concept of *g* may not be all that helpful toward explaining infants' performance on cognitive tasks, research on infant cognition may well contribute to an understanding of those cognitive processes that comprise *g*. We envision research on developmental and individual differences in infant processing skills as an important link toward establishing a unified developmental cognitive science. Such research will ultimately generate more powerful explanatory accounts of the origins of and/or the precursors to the processes under investigation in older children and adults. Ultimately, this will contribute to our understanding of the standardized assessment of those individual differences in overall cognitive function that we currently refer to as intelligence.

ACKNOWLEDGMENTS

Preparation of this chapter was supported in part by grant HD29960 from the National Institutes of Health to J.C. The authors thank Janet E. Frick, Rashmita Mistry, and the editors of this volume for their comments on earlier drafts.

NOTES

1. To wit:

> [Our] interest in intelligence...derives from a more general interest in the development and continuity of information processing from early in life...the real concern here is with individual differences in information processing, and what those individual differences imply for an understanding and measurement of information processing later in childhood and adulthood (Colombo, 1993, p. 3).

2. Although the developers of these infant tests have claimed that they were never meant to assess "intelligence," the titles used to label the infant tests, their format, their scoring, and their use suggest that it was assumed that these tests were designed to tap some aspect of mental or intellectual function early in the life span (see Colombo, 1993, pp. 11-14).

3. It should be noted that Binet himself had previously abandoned attempts to include infancy within the scope of his assessment instrument, although his co-worker Simon (1916) subsequently published a guide for assessing children under the age of 2 (see Reymert, 1920), and other adaptations of the Binet test were expanded to include items appropriate for infancy (see e.g., Kuhlmann, 1912).

4. The differential psychologists such as Nemzek (1930), Thorndike (1940), and Anastasi (1969) referred to stability in test performance across time as "constancy," and have stressed this characteristic as an expected characteristic of the measurement system. More contemporary developmentalists involved with the prediction-from-infancy phenomenon (including ourselves) have used the term "continuity," which seems (to us, at least) to imply reference more distinctly to the hypothesized underlying construct, rather than to its psychometric properties.

REFERENCES

Ackerman, P. L. (1986). Individual differences in information processing: An investigation of intellectual ability and task performance during practice. *Intelligence*, *10*, 101-139.

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology*, *117*, 288-318.

Anastasi, A. (1969). *Differential psychology*. New York: MacMillan.

Aslin, R. N., & Dumais, S. T. (1980). Binocular vision in infants: A review and a theoretical framework. In L. P. Lipsitt & H. W. Reese (Eds.), *Advances in child development and behavior* (Vol. 2). New York: Academic Press.

Bayley, N. (1933). *The California first-year mental scale*. Berkeley, CA: University of California Press.

Belmont, J. M., & Mitchell, D. W. (1987). The general strategies hypothesis as applied to cognitive theory in mental retardation. *Intelligence*, *11*, 91-105.

Benasich, A. A., & Bejar, I. I. (1992). The Fagan Test of Infant Intelligence: A critical review. *Journal of Applied Developmental Psychology*, *13*, 153-171.

Berg, C. A., & Sternberg, R. J. (1985). Response to novelty: Continuty vs. discontinuity in the developmental course of intelligence. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 15, pp. 1-47). New York: Academic Press.

Binet, A. (1903). *L'etude experimentale de l'intelligence* [The experimental study of intelligence]. Paris: Schleicher et Frenes.

Binet, A., & Henri, V. (1895). La psychologie individuelle [Individual psychology]. *L'Annee Psychologique*, *2*, 411-465.

Binet, A., & Simon, T. (1905). Sur la necessite d'etablir un diagnostic scientifique des etats de l'intelligence [On the necessity of establishing a scientific diagnosis of hte states of intelligence]. *L'Annee Psychologique*, *11*, 191-224.

Binet, A., & Simon, T. (1908). Le development de l'intelligence chez les enfants [The development of intelligence in infants]. *L'Annee Psychologique*, *14*, 1-94.

Binet, A., & Simon, T. (1911). *A method of measuring the development of the intelligence of young children*. Lincoln, IL: Courier.

Bornstein, M. H., & Sigman, M. D. (1986). Continuity in mental development from infancy. *Child Development*, *57*, 251-274.

Brooks-Gunn, J., & Weinraub, M. (1983). Origins of infant intelligence testing. In M. Lewis (Ed.), *Origins of intelligence* (2nd ed, pp. 25-66). New York, NY: Plenum.

Carroll, J. B., & Maxwell, S. E. (1979). Individual differences in cognitive abilities. *Annual Review of Psychology*, *30*, 603-640.

Cattell, J. (1890). Mental tests and measurements. *Mind*, *15*, 373-381.

Cattell, R. B. (1987). *Intelligence: Its structure, growth, and action*. Amsterdam, The Netherlands: Elsevier Science Publishers.

Ceci, S. L., Nightingale, N. N., & Baker, J. B. (1992). The ecologies of intelligence: Challenges to traditional views. In D. K. Detterman, (Ed.), *Current topics in human intelligence* [*Vol. 2*]*: Is the mind modular or unitary?* (pp. 61-82). New York: Ablex.

Colombo, J. (1993). *Infant cognition: Predicting childhood intelligence*. Newbury Park, CA: Sage.

Colombo, J. (1995). On the neural mechanisms underlying developmental and individual differences in infant fixation duration: Two hypotheses. *Developmental Review*, *15*, 97-135.

Colombo, J., & Mitchell, D. W. (1988). Infant visual habituation: In defense of an information-processing analysis. *European Bulletin of Cognitive Psychology/ Cahiers de Psychologie Cognitive*, *8*, 455-461.

Colombo, J., & Mitchell, D. W. (1990). Individual and developmental differences in infant visual attention. In J. Colombo & J. W. Fagen (Eds.), *Individual differences in infancy* (pp. 193-227). Hillsdale, NJ: Erlbaum.

DeLoache, J. S., Cassidy, D. J., & Brown, A. L. (1985). Precursors of mnemonic stragies in very young children's memory. *Child Development*, *56*, 125-137.

Detterman, D. K. (1992). *Current topics in human intelligence* [*Vol. 2*]*: Is the mind modular or unitary?*. Norwood, NJ: Ablex.

DiLalla, L. F., Thompson, L. A., Plomin, R., Phillips, K., Fagan, J. F., Haith, M. M., Cyphers, L. H., & Fulker, D. W. (1990). Infant predictors of preschool and adult IQ: A study of infant twins and their parents. *Developmental Psychology*, *26*, 759-769.

Eysenck, H. J. (1992). Intelligence: The one and the many. In D. K. Detterman (Ed.), *Current topics in human intelligence* (Vol. 2, pp. 83-116). Norwood, NJ: Ablex.

Fagan, J. F. (1971). Infant recognition memory for a series of visual stimuli. *Journal of Experimental Child Psychology*, *27*, 27-34.

Fagan, J. F. (1974). Infant recognition memory: The effects of length of familiarization and type of discrimination task. *Child Development*, *45*, 351-356.

Fagan, J. F. (1984a). The intelligent infant: Implications. *Intelligence*, *8*, 1-9.

Fagan, J. F. (1984b). The relationship of novelty preferences during infancy to later intelligence and recognition memory. *Intelligence*, *8*, 339-346.

Fagan, J. F., & Detterman, D. K. (1992). The Fagan Test of Infant Intelligence: A technical summary. *Journal of Applied Developmental Psychology*, *13*, 173-193.

Fagan, J. F., & Shepard, P. A. (1986/ 1987). *The Fagan test of infant intelligence*. Cleveland, OH: Infantest Corporation.

Fagan, J. F., Singer, J., Montie, J., & Shepard, P. A. (1986). Selective screening device for the early detection of normal or delayed cognitive development in infants at risk for later mental retardation. *Pediatrics*, *78*, 1021-1026.

Fagen, J. W., & Ohr, P. S. (1990). Individual differences in infant conditioning and memory. In J. Colombo & J. W. Fagen (Eds.), *Individual differences in infancy* (pp. 155-192). Hillsdale, NJ: Erlbaum.

Fillmore, E. A. (1936). *Iowa tests for young children*. Iowa City, IA: University of Iowa Press.

Flavell, J. H., & Wohlwill, J. F. (1969). Formal and functional aspects of cognitive development. In D. Elkind & J. H. Flavell (Eds.), *Studies in cognitive development: Essays in honor of Jean Piaget*. New York: Oxford University Press.

Furfey, P. H., & Muehlenbein, J. (1932). The validity of infant intelligence tests. *Journal of Genetic Psychology*, *40*, 219-233.

Galton, F. (1892). *Hereditary genius: An enquiry into its laws and consequences*. London, England: MacMillan.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gesell, A. L. (1925). *The mental growth of the preschool child*. New York: MacMillan.

Gottleib, G. (1976). Conceptions of prenatal development: Behavioral embryology. *Psychological Review*, *83*, 215-234.

Gottlieb, G. (1992). *Individual development and evolution: The genesis of novel behavior*. New York: Oxford University Press.

Gould, S. J. (1981). *The mismeasure of man*. New York: Norton and Company.

Gould, S. J. (1987). *An urchin in the storm*. New York: Norton and Company.

Guilford, J. P. (1967). *The nature of human intelligence*. New York, NY: McGraw-Hill.

Haith, M. M., & McCarty, M. E. (1990). Stability of visual expectations at 3.0 months of age. *Developmental Psychology*, *26*, 68-74.

Honzik, M. P. (1976). Value and limitations of infant tests: An overview. In M. Lewis (Ed.), *Origins of intelligence: Infancy and early childhood* (pp. 59-95). New York: Plenum Press.

Horn, J. L. (1985). Remodeling old models of intelligence. In B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurements, and applications*. New York: Wiley and Sons.

Horowitz, F. D., Paden, L. Y., Bhana, K., & Self, P. A. (1972). An infant control procedure for studying infant visual fixations. *Developmental Psychology*, *7*, 90.

Hunt, E. B. (1980). Intelligence as an information processing concept. *British Journal of Psychology*, *71*, 449-474.

Hunt, E. B. (1995). The role of intelligence in modern society. *American Scientist*, *83*, 356-368.

Jacobson, S. W., Jacobson, J. J., O'Neill, J. M., Padgett, R. J., Frankowski, J. J., & Bihun, J. T. (1992). Visual expectation and dimensions of infant information processing. *Child Development*, *63*, 711-724.

Jensen, A. R. (1979). *g*: Outmodeled theory or unconquered frontier? *Creative Science and Technology*, *2*, 16-29.

Jensen, A. R. (1984). Mental speed and levels of analysis. *Behavioral and Brain Sciences*, *7*, 295-296.

Jensen, A. R. (1993). Why is reaction time correlated with psychometric *g*? *Current Directions in Psychological Science*, *2*, 53-55.

Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. *Journal of Cognitive Neuroscience*, *3*, 335-344.

Keating, D. P. (1984). The emperor's new clothes: The "new look" in intelligence research. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. II, pp. 1-35). Hillsdale, NJ: Lawrence Erlbaum Associates.

Keating, D. P., & MacLean, D. J. (1987). Cognitive processing, cognitive ability, and development: A reconsideration. In P. A. Vernon (Ed.), *Speed of information-processing and intelligence* (pp. 239-270). Norwood, NJ: Ablex.

Kuhlmann, F. (1912). A revision of the Binet-Simon system for measuring the intelligence of children. *Journal of Psycho-Aesthetics*, *16*, 113-139.

Larson, G. E., & Alderton, D. L. (1992). The structure and capacity of thought: Some comments on the cognitive underpinnings of *g*. In D. K. Detterman (Ed.), *Current topics in human intelligence* (Vol. 2, pp. 141-156). Norwood, NJ: Ablex.

McCall, R. B. (1979a). Qualitative transitions in behavioral development in the first two years of life. In M. H. Bornstein & W. Kessen (Eds.), *Psychological development from infancy: Image to intention* (pp. 183-224). Hillsdale, NJ: Erlbaum.

McCall, R. B. (1979b). The development of intellectual functioning in infancy and the prediction of later IQ. In J. D. Osofsky (Ed.), *Handbook of infant development* (pp. 707-741). New York, NY: John Wiley and Sons.

McCall, R. B. (1983). A conceptual approach to early mental development. In M. Lewis (Ed.), *Origins of intelligence* (2nd ed., pp. 67-106). New York, NY: Plenum.

McCall, R. B. (1994). What process mediates prediction of childhood IQ from infant habituation and recognition memory? Speculations on the roles of inhibition and rate of information processing. *Intelligence. 18*, 107-124.

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

McCall, R. B., Eichorn, D. H., & Hogarty, P. S. (1977). Transitions in early mental development. *Monographs of the Society for Research in Child Development*, *42*(#3, Whole #171).

McCall, R. B., Hogarty, P. S., & Hurlburt, N. (1972). Transiitions in infant sensorimotor development and the prediction of childhood IQ. *American Psychologist*, *27*, 728-748.

Miller, D., Spiridigliozzi, G., Ryan, E., Callan, M., & McLaughlin, J. (1980). Habituation and cognitive performance: Relationships between measures at four years of age and earlier assessments. *International Journal of Behavioral Development*, *3*, 131-146.

Nemzek, C. L. (1933). The constancy of IQ. *Psychological Bulletin*, *30*, 143-167.

O'Connor, M., Cohen, S., & Parmelee, A. (1984). Infant auditory discrimination in preterm and fullterm infants as a predictor of 5 year intelligence. *Developmental Psychology*, *20*, 159-165.

Pellegrino, J. W., & Glaser, R. (1979). Cognitive correlates and components in the analysis of individual differences. *Intelligence*, *3*, 187-214.

Reed, T. E., & Jensen, A. R. (1992). Conduction velocity in a brain nerve pathway of normal adults correlates with intelligence. *Intelligence*, *16*, 259-272.

Resnick, L. B. (1976). *The nature of intelligence*. Hillsdale: Lawrence Erbaum Associates.

Reymert, M. L. (1920). Questionnaire for the observation of a young child from birth to two months of age. *Pedagogical Seminary and Journal of Genetic Psychology*, *27*, 200-204.

Rose, D., Slater, A., & Perry, H. (1986). Prediction of childhood intelligence from habituation in early infancy. *Intelligence*, *10*, 251-263.

Rose, S. A. (1989). Measuring infant intelligence: New perspectives. In M. H. Bornstein & N. A. Krasnegor (Ed.), *Stability and continuity in mental development: Behavioral and biological perspectives* (pp. 171-188). Hillsdale, NJ: Erlbaum.

Rose, S. A., & Feldman, J. F. (1995). Prediction of IQ and specific cognitive abilities at 11 years from infancy measures. *Developmental Psychology*, *31*, 685-696.

Rose, S. A., Feldman, J. F., & Wallace, I. F. (1992). Infant information processing in relation to six-year outcomes. *Child Development*, *63*, 1126-1141.

Rose, S. A., Feldman, J. F., Wallace, I. F., & Cohen, P. (1991). Language: A partial link between infant attention and later intelligence. *Developmental Psychology*, *27*, 798-805.

Rose, S. A., Feldman, J. F., Wallace, I. F., & McCarton, C. M. (1989). Infant visual attention: Relation to birth status and developmental outcome during the first 5 years. *Developmental Psychology*, *25*, 560-576.

Rose, S. A., & Wallace, I. F., & Feldman, J. F. (1988). Individual differences in infant informtion processing: Reliability, stability, and prediction. *Child Development*, *59*, 1177-1197.

Sigel, I. E. (1963). How intelligence test limit understanding of intelligence. *The Merrill-Palmer Quarterly*, *9*, 39-56.

Shirley, M. M. (1933). *The first two years: A study of twenty-five babies. II. Intellectual development*. Minneapolis, MN: University of Minnesota Press.

Sigman, M. D., Cohen, S. E., Beckwith, L., Asarnow, R., & Parmelee, A. H. (1991). Continuity in cognitive abilities from infancy to 12 years of age. *Cognitive Development*, *6*, 47-57.

Simon, T. (1916). Les deux premieres annees de l'enfant [The first two years of the infant]. *Bulletin de la Societe libre Etude psychologie des Enfants*, *108*, 1-64.

Snow, R. E. (1982). The training of intellectual aptitude. In D. K. Detterman & R. J. Sternberg (Eds.), *How and how much can intelligence be increased*. Norwood, NJ: Ablex.

Spearman, C. (1904). General "intelligence" objectively determined and measured. *American Journal of Psychology*, *15*, 201-293.

Spearman, C., & Jones, H. (1950). *Human ability*. London, England: MacMillan.

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

Sternberg, R. J. (1985b). *Human abilities: An information processing approach*. San Francisco, CA: W. H. Freeman.

Stott, L. H., & Ball, R. S. (1965). Infant and preschool mental tests: Review and evaluation. *Monographs of the Society for Research in Child Development*, *30*(#3, Whole #101).

Stutsman, R. (1931). *Mental measurement of preschool children with a guide for the administration of the Merrill-Palmer scale of mental tests*. Yonkers-On-Hudson, NY: World Book.

Terman, L. M. (1916). *The measurement of intelligence*. Boston, MA: Houghton-Mifflin.

Thorndike, R. L. (1940). "Constancy" of the IQ. *Psychological Bulletin*, *37*, 167-186.

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

Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago, IL: University of Chicago Press.

Thurstone, L. L., & Thurstone, T. G. (1941). *Factorial studies of intelligence*. Chicago, IL: University of Chicago Press.

Tyler, L. E. (1976). The intelligence we test: An evolving concept. In L. B. Resnick's (Ed.), *The nature of intelligence*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *American Psychologist*, *30*, 128-134.

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

Vernon, P. A., & Mori, M. (1992). Intelligence, reaction times, and peripheral nerve conduction velocity. *Intelligence*, *16*, 273-288.

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

Yerkes, R. M. (1921). Psychological examining in the United States army. *Memoirs of the National Academy of Sciences*, *15*, 1-89.

INTELLIGENCE:

THE NEW LOOK

H.J. Eysenck

THE SEARCH FOR A PARADIGM

Recent developments in the study of intelligence can best be understood in terms of Kuhn's (1957, 1970) theory of normal science and scientific revolutions, a theory which applies to social science just as much as to physical science (Barnes, 1982). Normal science operates according to a universally accepted paradigm. The term was used very loosely in Kuhn's original writings (Masterman, 1970), and Kuhn (1974) has since redefined it more cogently. In the context of this paper, the term "paradigm" will be used in the sense of "exemplar," that is, an accepted problem-solution in science, related to a particular concrete scientific achievement. Such a paradigm has gained universal acceptance throughout the scientific field as a valid procedure, and as a model of valid procedure for pedagogic use. Students are taught the paradigm; they work with it; and they try to solve the problems created by failure of experiments to agree with predictions made in terms of the paradigm.

Advances in Cognition and Educational Practice, Volume 4, pages 119-148.

ISBN: 0-7623-0105-8

However, some anomalies arise which persist in spite of efforts to accommodate them to the paradigm; they defy analysis as a familiar in unfamiliar form. The puzzle-solving activities of normal science, which attempt to show that what is prima facie anomalous is either the spurious product of bad equipment, of technique, or a familiar phenomenon in disguise, may fail, and over time such recalcitrant anomalies may accumulate and generate a feeling that all is not well with the paradigm. This is a crisis situation which may and often does lead to a scientific revolution.

Crises of this kind typically involve changes in the character of research. Speculation may become more acceptable; novel and radically deviant procedures and interpretations may be tried; and the paradigms, and the activities and judgments based upon them may be called into question. I am suggesting that such a crisis has occurred in the field of intelligence measurement, that the widespread suspicion that something is at fault at the heart of normal research itself has led to changes in the character of research, involving speculation of a novel kind, new and radically deviant procedures and interpretations, and the development of new theories altogether.

Actually the course of revolution described by Kuhn is only one of several possible models. We may indeed deal with a paradigm that generates too many false predictions, and has to be supplanted by a new paradigm. However, an alternative might be that such a questionable paradigm is supplanted by a paradigm already familiar and previously discarded. The Copernican revolution might be such an example; the heliocentric theory of Copernicus had already been advanced two thousand years before by Aristarchus, only to be displaced by the Ptolemaic geocentric paradigm; Copernicus was in essence advocating a return to a previous paradigm. The same would be true of the wave theory of light; when Thomas Young and the French physicists overthrew Newton's corpuscular paradigm they were in fact returning to a previously existing and earlier wave paradigm. It is this alternative kind of paradigm-change which appears to be taking place in psychology, where the paradigm that held sway from 1910 until recently was preceded by another paradigm which is now returning as a result of the most recent revolution (Eysenck, 1983).

The original paradigm in question was that of Galton (1879), and that which replaced it was that of Binet (Wolf, 1973). The two paradigms differ in vital respects, and while the Binet-type paradigm was widely accepted during the past 70 years or so, and as far as measurement is concerned might almost be considered to have been unrivalled, recent work has suggested that in fact the Galton paradigm was superior, and that it can explain anomalies with which the Binet-type paradigm cannot deal at all. Table 1 sets out the major differences between Galton and Binet.

Table 1. Main Differences Between the Galton and Binet Paradigm on Intelligence

Difference No.	Paradigm	
	Binet	Galton
1.	Concept of intelligence is the average of a number of independent abilities.	Concept of intelligence is a unitary cognitive ability underlying all mental performance.
2.	Focus on environmental modifications of phenotypic behavior.	Focus on inherited dispositional factor determining differences in phenotypic performance.
3.	Measurement of intelligence by means of test items mirroring different aspects of intelligence.	Measurement of intelligence by means of physiological types of tests.

The revolution which is taking place now is also interesting from the point of view of another observation made by Kuhn. Revolutions, he says, constitute discontinuities in research and the growth of knowledge; learned similarity relations are modified; concrete instances are being grouped into new clusters; and the whole conceptual fabric is reconstructed so that concepts figure in different laws and generalizations. Definitions of such terms as "compound" and "mass" are different prior to and after the scientific revolutions produced by Dalton and Einstein, and, as Barnes (1982) points out, "The reconstruction of practice results in new modes of cognition, inference and explanation." Thus we may find that the very definition of "intelligence" in the two paradigms, that is, that of Binet and that of Galton, may be different, and that the two paradigms may indeed be incommensurable.

These different definitions of "intelligence" have for a long time bedevilled all discussions, both scientific and popular, about the measurement of intelligence, the heritability of intelligence, and the practical applicability of tests of intelligence. Essentially, we must distinguish between intelligence A, meaning the biological and genetically determined basis for all cognitive performance, and all individual differences in cognitive performance; intelligence B, which is intellectual ability shown in everyday life, and thus adulterated by differences in personality, education, socio-economic status, cultural factors, and the thousand and one accidents of life which may affect a person's acquisition of knowledge, problem-solving capacity, and so on. Finally, intelligence C is the intelligence which we measure in various ways, such as by IQ tests or by means of such psychophysiological techniques as will be described later on in this paper. Figure 1 shows in brief outline the relationship between intelligence A, intelligence B, and intelligence C.

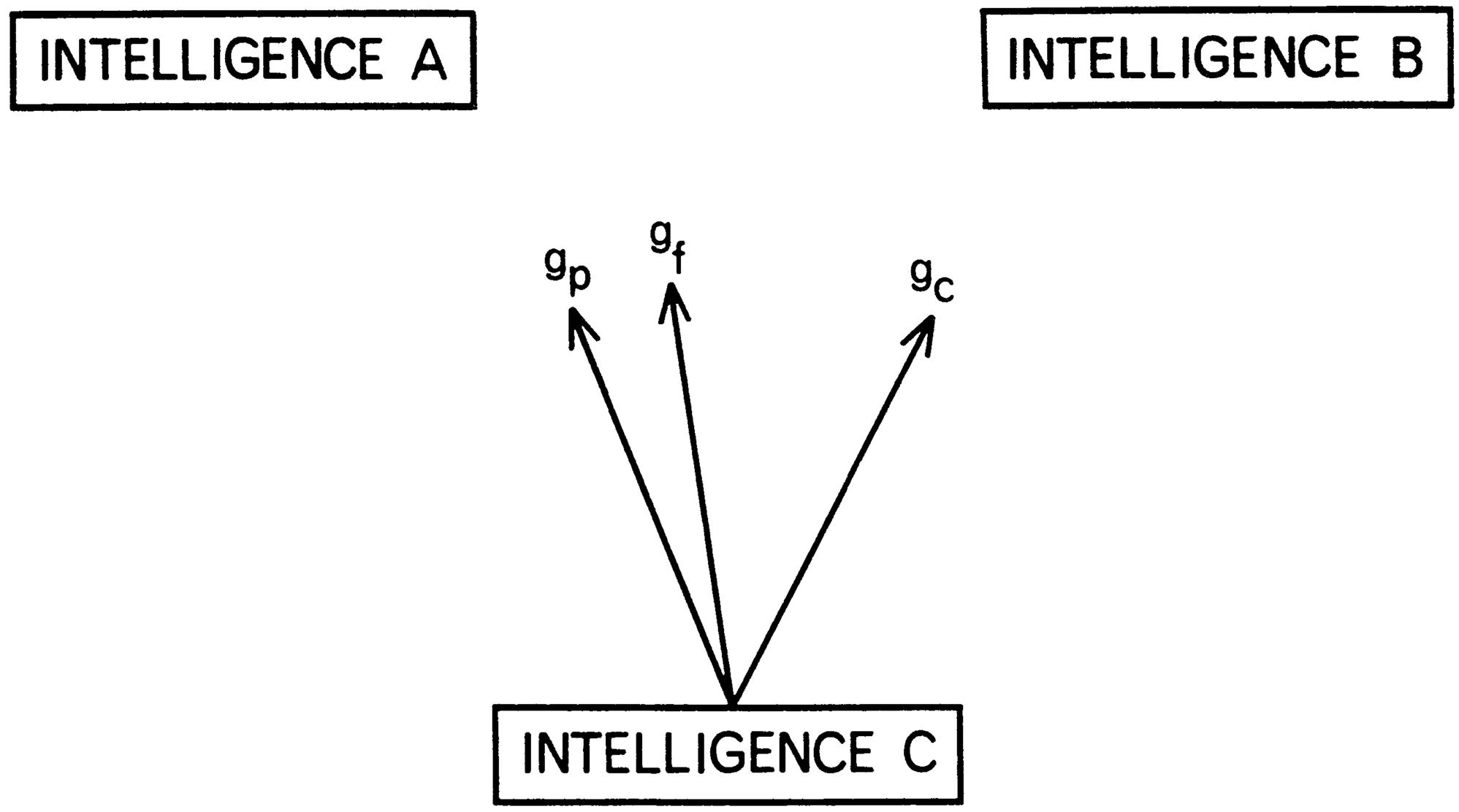


Figure 1. Intelligence A, B and C, Related to g_c (Crystallized Ability) g_f (Fluid Ability), and g_p (Psychophysiological Basis of Intelligence)

In Figure 1, three different types of intelligence measures are distinguished. To measure the general factor of intelligence (g) we may use tests of what Cattell (1963, 1971) calls fluid ability (g_f), that is, relatively culture-fair tests which do not call upon any great degree of acquired knowledge, linguistic skill and so on, or tests of what he calls crystallized ability (g_c), which in practice means verbal tests of IQ relying to a greater or lesser extent on acquired knowledge of various kinds. Finally, we may deal with physiological or semiphysiological (reaction time) tests, symbolized as g_p in Figure 1, that is, measures of physiological structure and functioning underlying differences in mental ability. Clearly, as the arrows show, tests of g_p are closer to intelligence A, tests of g_c closer to intelligence B, and tests of g_f somewhere inbetween the other two.

The differentiation between intelligence A and intelligence B is crucial for an understanding of the two paradigms to which we will turn next; it is also crucial for an understanding of the wide differences in attitudes towards the measurement of intelligence, and the acceptability of its usefulness in scientific discourse, which has figured so prominently in debates over the past 50 years or so. It has sometimes been suggested that intelligence A cannot accommodate the rich diversity of factors encompassed by intelligence B and it has equally been suggested that intelligence B is too complex to be a properly measurable and acceptable scientific concept. Both objections are reasonable as long as we use a single term, namely "intelligence," without discriminating, by subscript or otherwise, between the different conceptions. It should be noted, however, that this situation is not unique in science; we find exactly the same in relation

to such accepted scientific concepts as "heat," and the associated practice of temperature measurement. The kinetic theory of heat gives us a perfectly simple and intelligible definition of heat in terms of the speed of movement of molecules; this might be called heat A. However, there is also a heat B, that is, the existing temperature as individually perceived by people in their everyday lives.

Heat B is of course determined to a large extent by heat A, but it is also determined in part by factors which are irrelevant to heat A, such as humidity, percentage of fat in the body, food eaten and particularly alcohol intake during the day, the chill factor (i.e., the amount of wind present), radiant heat, and so on. An experimental demonstration of the difference between heat A and heat B is given in the well known experiment where the subject, who is blindfolded, immerses his right hand in a bowl of hot water, and his left hand in a bowl of cold water. After a few minutes he immerses both his hands in a bowl of lukewarm water, which feels cold to his right hand, but hot to his left hand. Thus while the actual heat A, as measured, is equal for the two hands, heat B differs widely from one to the other.

Figure 2 illustrated the nature of these three "intelligences" in some greater detail. Biological intelligence is fundamental to all cognitive activity; it is largely determined by heredity, and is measurable by electrophysiological indices, such as the averaged evoked potential (AEP), the galvanic skin response (GSR), the electroencephalograph (EEG), the contingent negative variation (CNV), positron emission tomography (PET scan), magnetic resonance imaging (MRI), and many others to be discussed in a later section. IQ measures psychometric intelligence, produced in part by biological factors,but also by environmental ones. And IQ finally gives rise to social or practical intelligence, which in turn is determined in large part by additional environmental factors. None of these is right or wrong; they simply reflect different ways of looking at intelligence (Eysenck, 1988).

INTELLIGENCE: THE TWO PARADIGMS

The paradigms of intelligence developed by Galton and Binet differ on three major points. The first of these concerns the very concept of intelligence itself. For Galton this was a central, unitary cognitive ability, underlying and influencing all mental performance, and basic to all types of problem-solving. Binet, on the other hand, although he used the term "intelligence," really defined it out of existence by stating that it was essentially the average of a number of disparate and probably independent abilities, all of which had to be measured by his newly developed tests. Thus right from the beginning Binet considered intelligence to be a relatively arbitrary concept based on the much more firmly established and "real" primary abilities which he postulated.

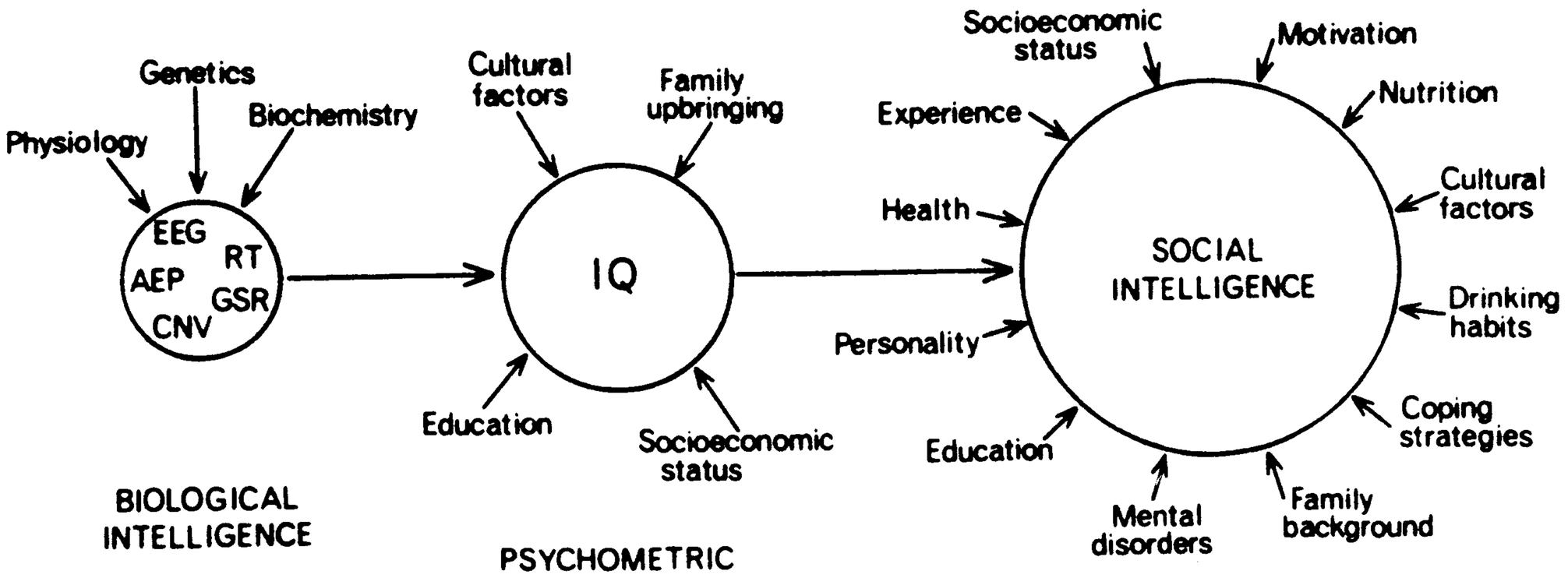


Figure 2. Sequence From Biological Intelligence Through Psychometric to Social or Applied Intelligence

This difference is sometimes erroneously identified with the question of *reification*. Thus it is sometimes said that Galton "reified" the concept of intelligence, by conceiving of it as really existing, whereas Binet, by considering it a statistical and conceptual artefact, escaped the accusation of reification. This argument is strictly inadmissible. Intelligence, as I have pointed out elsewhere (Eysenck, 1979) is a scientific concept, and as such can claim no real "existence" in the sense that a table or a pig, or a book can claim existence. *All concepts are inventions of the human mind, and intelligence is no different from such concepts as heat, velocity, or gravitation.* Thus Galton's conception implies no reification, and his conception of intelligence is no more artificial than is Binet's. Some writers, arguing on both sides, have indeed used expressions which implied reification, i.e. by arguing about the "existence" of intelligence, but reification was probably not intended by any of them. It is difficult to avoid the use of terms which seem to imply reification, but the danger of misunderstanding is probably less than is sometimes believed.

The question raised by Galton and Binet is still important. Spearman (1927) introduced the statistical method of factor analysis into the field in order to settle the issue, and emerged with his general factor of intelligence (*g*), but Thurstone (1938) failed to find a general factor, and postulated instead a number of independent primary mental abilities. Eysenck (1939) reanalysed Thurstone's matrix of intercorrelations and found evidence both for a general factor and for a number of group factors closely resembling Thurstone's primary mental abilities. A number of faults in the experimental design, such as the use of subjects showing a very restricted range of intelligence, vitiated Thurstone's original experiment, and on repetition using a wider range of

ability in his subject pool Thurstone and Thurstone (1941) arrived at the same conclusion as Eysenck (1939), namely that the data demanded both a general factor of intelligence, and also a number of primary mental abilities, a conclusion which Spearman later assented to (Eysenck, 1992).

Most psychometrists nowadays would probably accept such a general conclusion (Eysenck, 1979; Vernon, 1979; Brody, 1992), but Guilford (Guilford, 1967; Guilford & Hoepfner, 1971) still advocates a "structure-of-intellect" model of intelligence which transforms the Binet system into a reductio ad absurdum. Figure 3 shows this model, which postulates five different operations on four different types of content, producing six different types of product; all possible combinations of these three classes give rise to 120 different abilities which are conceived of as being quite independent of each other. Psychometric evidence produced by Guilford fails to establish the viability of this model, and there are many reasons for rejecting it as being completely unrealistic (Eysenck, 1979). Nevertheless, in the last resort the distribution of variance over different tests is to some extent subjective and arbitrary, and it is certainly possible to distribute the variance in such a way as to describe the test performance of Guilford's subjects in terms of 120 independent variables. This may indicate that the argument is not really resolvable in terms of correlational and factor analytic techniques only, and that recourse must be had to experimental studies based on specific theories about the nature of *g*. Carroll (1993) gives a very complete account of the present state of the factorial analysis of intelligence.

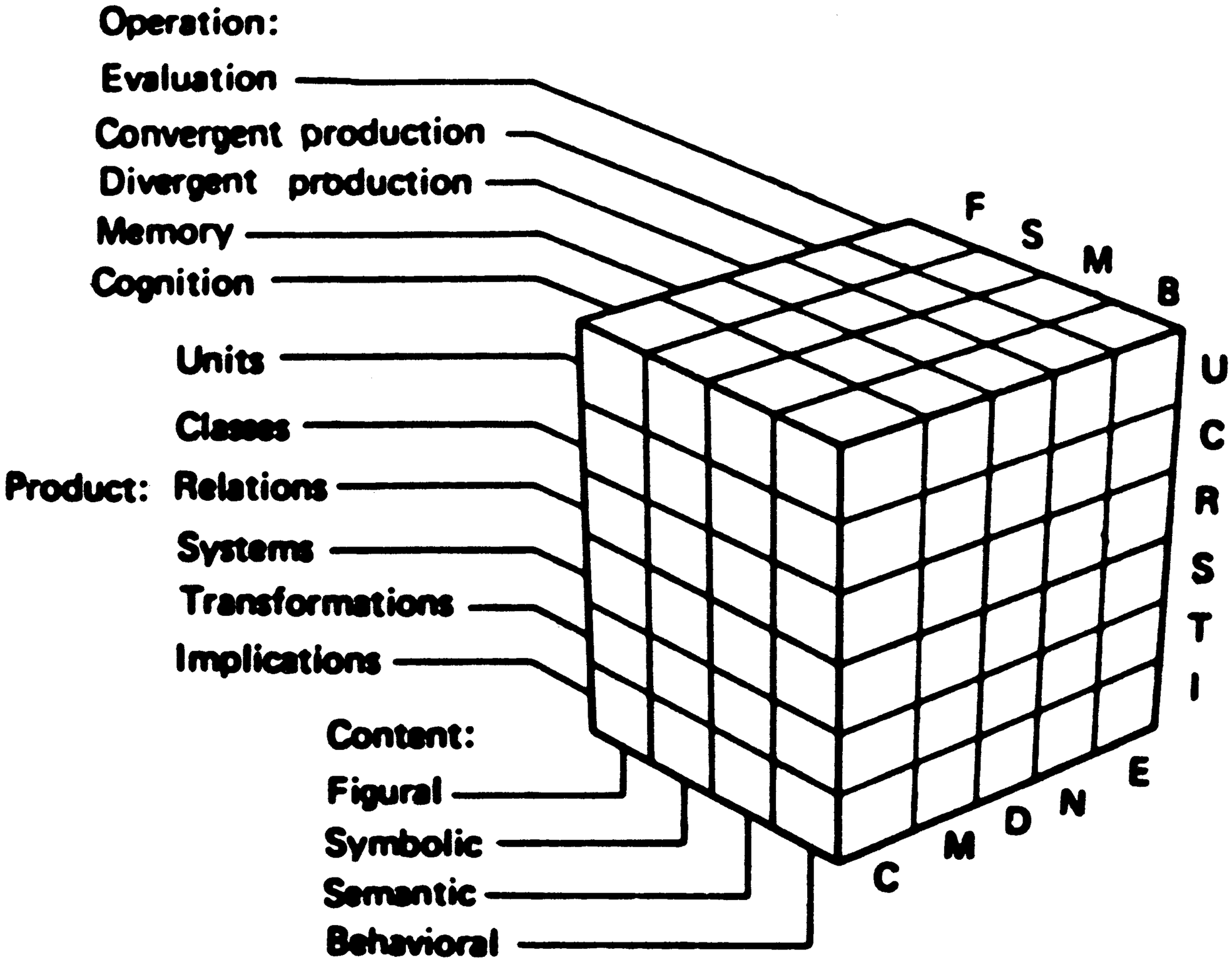


Figure 3. Guilford's Structure-of-Intellect Model of Intelligence

The second great division between Galton and Binet relates to heritability. Galton assumed that intelligence was essentially an inherited dispositional factor determining differences in phenotypic performance. Binet, while not denying the possibility of genetic influences, was as an educationalist far more interested in environmental modifications of phenotypic behavior, and hence set the scene for the debate between hereditarians and environmentalists which has so disfigured the scientific scene in recent years.

The evidence on this point has been reviewed in great detail elsewhere (Eysenck, 1979), and the outcome appears to be that IQ measures of the usual kind, such as the Wechsler or the Binet, show genetic contributions to the variance of something like 80 percent, and of environmental factors of something like 20 percent, with little evidence for interaction factors. (This conclusion is based on many different types of studies, but excluding the work of Burt which has been impugned by critics as being unreliable.) It should be noted that other writers (e.g., Vernon, 1979) suggest rather lower heritabilities, but two points are important here. The first one is that, like intelligence, so also heritability is a concept that has to be carefully defined before it can be used scientifically. At least three different uses have been noted. We may define heritability in terms of the additive genetic variance divided by the phenotypic variance; this is the so-called *narrow* heritability. The second meaning of heritability is given by the use of the total genetic variance, that is, incorporating dominance factors, assortative mating factors, and epistasis in addition to additive genetic variance, divided by the phenotypic variance; this is the so-called *broad* heritability. As dominance and assortative mating are certainly, and epistasis possibly, factors influencing genetic variance, the broad heritability is significantly greater for IQ differences than is the narrow heritability. Both formulae incorporate error variance in the environmental part of the equation, and this should be corrected for; such a correction for attenuation significantly raises the heritability, whether narrow or broad.

We may thus from one single study derive three heritabilities, such that the narrow heritability, uncorrected, might be something like 50 percent, the broad heritability, uncorrected might be 75 percent, and the broad heritability, corrected for attenuation, might be 80 percent. Different heritabilities would be quoted and used for different purposes, and the impression might be given to the uninitiated that quite different results have been achieved. The heritability of 80 percent mentioned above relates to the broad heritability, corrected for attenuation. Uncorrected estimates, or estimates of narrow heritability, would obviously be much lower. We must also bear in mind that heritability estimates are population estimates, which will differ from one population to another; there is no reason why American studies and English studies should give identical figures, or why studies carried out now should give identical figures to those carried out 50 years ago. Finally we should bear in mind that heritability estimates have a relatively large standard error

attached to them, unless the number of twins, adopted children, and so on, used is very large. As a consequence, estimates from different studies may differ to quite a large extent, without throwing doubt on the high degree of heritability of IQ. Thus on this, as on the previous point, I would suggest that Galton had the better of the argument, but that insofar as IQ tests measure intelligence B rather than intelligence A, both the existence of primary mental abilities and the importance of environmental factors, as postulated by Binet, have to be borne in mind.

The fact that heritability coefficients are population estimates applies not only to national and racial groups, but also to age groups. Thus heritability of the IQ is much higher for adults than for children, where it might only be 40 percent or 50 percent (Plomin & McCleary, 1993). The reason for this difference is presumably the fact that the cognitive behavior of children is much more narrowly circumscribed than that of adults, who are free to read or not to read; parents and teachers profoundly affect such choices for children.

The third great difference between Galton and Binet relates to the actual *measurement* of the different kinds of intelligence postulated by these two authors. Galton, as one might have expected, wanted intelligence to be measured by means of *physiological* types of tests, and in the absence of anything more direct he suggested measures of reaction time. Binet, on the other hand, and in accord with his definition of intelligence suggested test items which would mirror the kind of behavior which would be recognizable as instances of intelligence B, that is, problem-solving, obeying instructions, learning, memorising and reproducing information, both short-term and long-term, and generally demonstrating the ability to use information. As is well known, practically all IQ tests used since Binet's day have borne the stamp of his theory, and have incorporated test items of the kind he advocated. Culture-fair tests, such as those of Raven and Cattell, may seem to be an exception to this rule, but they too are subject to environmental influences, learning and the development of strategies, and while this influence of cultural factors may be less marked than in the case of tests of crystallized ability, such influences are by no means absent.

Some interesting questions are raised by the historical developments of the past seventy years or so. Why was the Binet-type of test accepted so readily, while the Galton-type of test was not only rejected, but was never seriously tried? What was the influence of the Zeitgeist on these developments, and what was the influence of serious scientific reasoning and experimentation? Has the Binet paradigm been sufficiently successful to explain the anomalies which inevitably occur in the development of any scientific theory? As we have seen, the evidence suggests that on two out of the three points, at least, Galton had the better of the argument; there is strong evidence of powerful genetic factors, and it is difficult to deny the (psychometric) existence of a general factor of intelligence, underlying all types of cognitive ability. If on these two points

Galton was right, rather than Binet, is it not possible on the third point also Galton might have been right (Eysenck, 1993)?

THE CHRONOMETRY OF INTELLIGENCE

Galton, in the absence of any truly psychophysiological methods of measurement existing at the time, had suggested simple behavioral indices, such as reaction time (RT), and physical measures, such as head size as being likely measures of intelligence. Many early attempts to relate head size and intelligence gave quite low but always positive correlations; more recently non-intrusive measures of *brain size* (only very poorly reflected in *head size*) have given much more positive results, in the neighborhood of 0.50 (Jensen & Sinha, 1993; Andreasen et al., 1993; Jensen & Johnson, 1994). There is now no doubt that through the use of magnetic resonance imaging we can demonstrate that Galton was right in his hypothesis.

Galton's theory of a relationship between intelligence and reaction time (RT) was subjected to the test by Wissler (1901). He was the first to use a coefficient of correlation to measure the degree of relationship between simple RT and "intelligence" as indexed by the course grades obtained by men students in Columbia College. The correlation turned out to be quite insignificant, but the experiment was singularly deficient from the point of view of methodology. Each subject's RT was based on an average of only 3-5 measurements,which we now know would result in an exceedingly low reliability; we would need something like 50 or more measurements in order to get a stable average for any given individual. The "range of talent" was highly restricted in this very selected group of university students, and of course such restriction greatly attenuates correlations between any *g*-loaded measurements. Finally, the reliability and validity of course grades as a measure of intelligence leave much to be desired, the best present-day IQ tests showing quite low correlations with grades in highly selective colleges.

In spite of its very poor quality, Wissler's work has since been quoted in many textbooks, and officially marked the end of the quest for information on the Galton hypothesis. Occasional studies, such as those by Peake and Boring (1926) and Lemmon (1927) did find positive correlations, but as Jensen (1982b) has pointed out, correlations between RT test scores were "an anathema to the psychometric Zeitgeist, which was much more bent on developing tests with practical predictive validity than on experimenting with laboratory techniques to investigate the nature of intelligence" (p. 98).

Lemmon's work had already suggested that choice RT may be more highly correlated with IQ than simple RT, but the development of this insight had to wait for the discovery of Hick's law (Hick, 1952). This law states that multiple choice RT increases as a linear function of the increase in amount of

information in the stimulus array, when information is measured in *bits*, that is, the logarithm to the base 2 of the number of choices. Roth (1964), using multiple choice RTs in an experimental paradigm based on Hick's law, found that individual differences in the *slope* of RT as a function of bits (i.e., the rate of information processing) were correlated with IQ. On the theory that an IQ test measures (among other things) information processing capacity, such a relationship between RT and intelligence can be predicted. The differences in the amounts of knowledge and skills called for by ordinary IQ tests differentiate between individuals because these differ in the rates with which they process (and hence "acquire") the information offered by the environment. As Jensen (1982b) says, "Other things being equal, individuals with greater speed of information processing acquire more cognitively integrated knowledge and skill per unit of time that they interact with the environment. Seemingly small individual differences in speed of information processing, amounting to only a few milliseconds per *bit* of information, when multiplied by months or years of interaction with the environment can account in part for the relatively large differences observed between individuals in vocabulary, general information, and the other developed cognitive skills assessed by IQ tests" (p. 99). Eysenck (1967) drew attention to the importance of these data, and their relevance to a theory of intelligence, and since this time a great deal of work has been done in connection with this paradigm. Jensen (1982a, 1982b) gives a good account of these developments, and the general conclusions that may be drawn from them.

Figure 4 shows the subjects' console often used for RT investigations. The subject puts his index figure on the central "home" button. When one of the green jewelled lights on the circumference lights up, S transfers his finger from the "home" button to the response button in front of the light in question. We measure *decision time* (DT), from the moment the target light appears to the moment the finger is lifted from the "home" button, and movement time (MT), from the moment the finger is lifted from the "home" button to the moment the target button is depressed. DT + MT = RT. The measure may be of simple RT (one light), choice RT (one of several lights), or complex RT, as for instance when 3 lights come on simultaneously, two close together, one (the odd man out) at some distance from the other two. Response has to be made to the odd man out (Frearson & Eysenck, 1986).

There would be little point in here citing all the studies leading to the conclusion that Galton (1883) and Sully (1876) were right in their proposal to use reaction time as a measure of intelligence. The work of such men as Baumeister and Kellas (1968), Brand and Deary (1982), Jensen and Munro (1974), Jensen et al. (1981), Keating and Bobbitt (1978), Lally and Nettlebeck (1977), and Vernon (1981) establishes a number of facts which are no longer in doubt. The major facts are these.

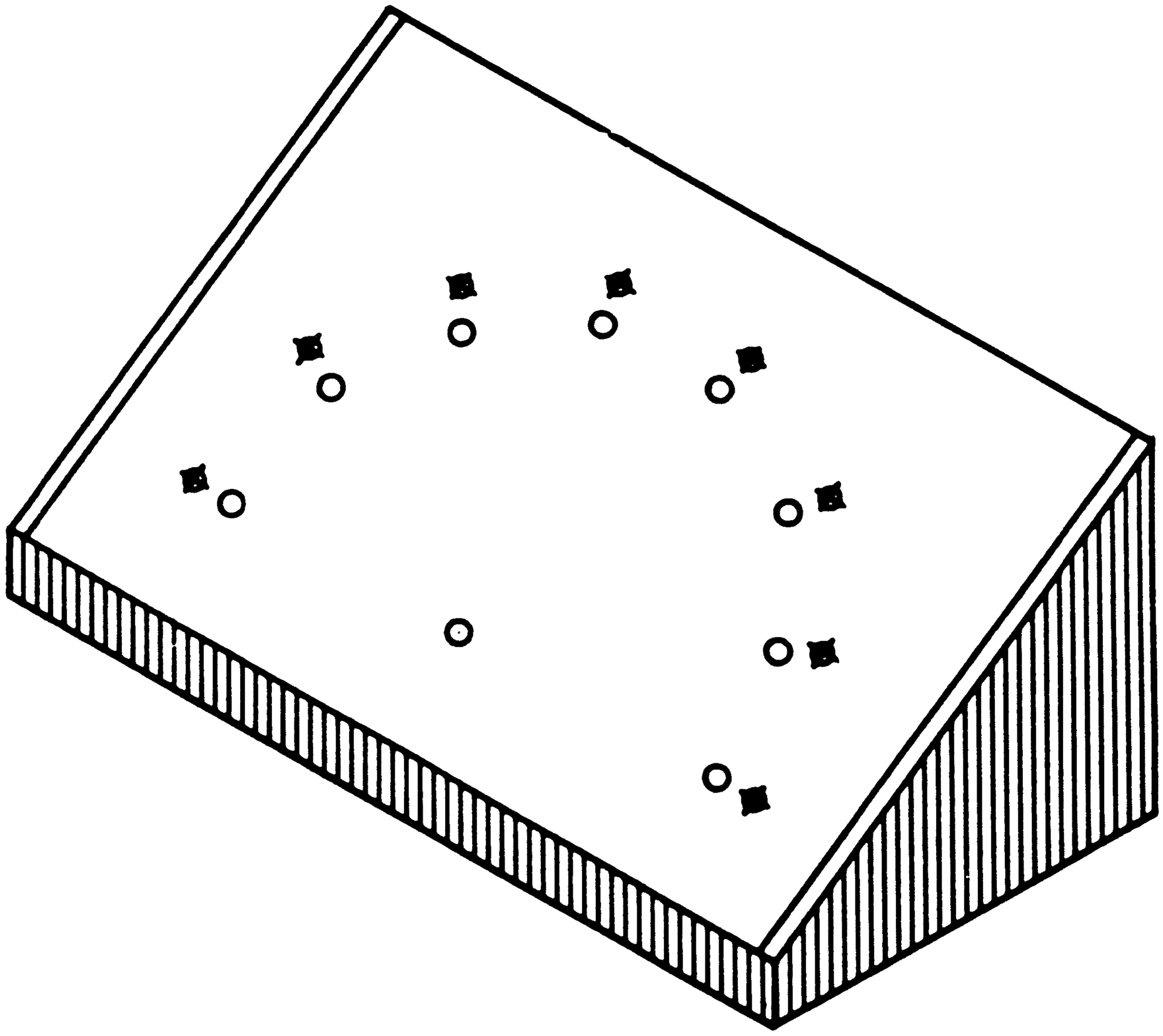


Figure 4. Subject's Console for Reaction Time Experiments

1. Simple reaction time correlates negatively with intelligence, as measured by orthodox IQ tests, such as the Wechsler, the Binet, Raven's Matrices, and so forth.
2. Choice reaction time correlates to a significantly higher degree with IQ than does simple reaction time.
3. Up to about four bits, the correlation of choice reaction time with IQ is greater, the larger the number of bits involved.
4. The slope relating reaction time to number of bits, according to Hick's law, correlates negatively with IQ, i.e. the steeper the increase in reaction time with increase in number of bits, the lower the IQ of the person concerned.
5. The variability of RT, whether in simple or choice reaction time experiments, is inversely related to the IQ of the subject; the higher the IQ, the lower the variability.

It will be noted that no specific correlation coefficients have been given in the above summary. The reason for this is that much of the work has involved comparisons between groups (students, persons of average ability, retardates), and it would not be possible to derive meaningful correlations from such comparisons. Another reason is that where correlations have been given, the groups tested were often unusually low or unusually high as far as the range of ability was concerned, thus decreasing or increasing the true correlation beyond what it should be. However, it should be noted that whatever differences between groups may be observed or whatever correlations are calculated, these are usually derived from experiments on one given day. It is well known that there are considerable diurnal variations in RT measure for any given subject, so that the observed correlations are minimal; for theoretical purposes they should be corrected for attenuation.

As an estimate, I would say on the basis of the published evidence that correlations between IQ and simple RT would probably be around -0.1; correlations with choice RT between -0.2 and -0.3. Correlations between IQ and variability would be around -0.4. Correlations with complex tests like the odd-man-out have been as high as -0.6. Combining these different measures in a multiple *r* would give us correlations approximating 0.7 when corrected for attenuation. These correlations are estimates for a randomly selected group of subjects, and of course the values are very approximate; nevertheless, I think they capture the rough range of values found in many different investigations. We may thus say that there is a surprisingly close relationship between reaction time measures and IQ, much closer than would have been expected on the basis of the Binet-type cognitive theories of intelligence to be found in such books as Sternberg's (1982 *Handbook of Human Intelligence*).

Of considerable relevance is the fact that simple RT shows a fairly linear decrease with age between about 5 and 15 years of age. Thus RT, on Binet's own showing, would be a good candidate for inclusion in a set of tests selected on the basis of showing linear development with age. Jensen (1982b) has also shown that the slope of regression of mean RT on age increases markedly as a function of the bits of information conveyed by the RS. This indicates that performance on the more complex RT tasks, involving a greater number of bits, reflects age differences much more sharply than does performance on simpler RT tasks. These results agree well with those discussed previously.

As the addition of choice reaction time paradigms and the Hick's slope constituted an enormous improvement on the simple RT type of experiment, so it was hoped that the elaboration of more cognitive types of experiments involving short-term and long-term memory would increase the relationship between IQ and RT. One such paradigm is the Sternberg (1966) test, which measures the subject's speed of scanning his short-term memory for information. The subject is shown a series of between two to seven digits or letters (termed the "positive set"), for several seconds, after which a single "probe" digit or letter

is presented. In a random half of the trial the probe digit is a member of the positive set, and the subject is required to respond as quickly as possible to the probe digit by pressing either a "yes" or a "no" button to indicate whether the probe was or was not a member of the positive set.

Several studies (e.g., McCauley et al., 1976; Keating & Bobbit, 1978) have shown a relationship between the intercept and slope of RT as a function of set size and mental test scores. These are impressive, but the involvement of short-term memory scan in the RT paradigm does not seem to have increased the correlation with IQ to any marked extent.

The Posner long-term memory access paradigm (Posner, 1969) is a measure of time it takes the subject to access a highly overlearned item of information stored in his long-term memory. The experimental procedure is based on the comparison by a subject of pairs of stimuli which are the same or different either physically or semantically; the score is the RT from exposure to answer. Hunt (1976) and Keating and Bobbitt (1978) reported correlations in the expected direction with intelligence, but again the hope for increase in size over choice RTs were not found.

A third paradigm involving more cognitive content than the ordinary RT experiment is the so-called "inspection time paradigm" of Lally and Nettlebeck (1977). This consists of the establishment of an inspection time threshold for the discrimination of two lines of distinctly unequal length, presented tachistoscopically; the shorter the inspection time needed, the higher is the individual's IQ. Actually experiments of a similar kind were already reported by Burt (1909), who indicated correlations in excess of .80; it has never been clear why he abandoned this very promising avenue! More recently Lehrl et al. (1980) has based a very promising intelligence test on what is essentially an inspection time paradigm. Brand and Deary (1982) discuss work on the relationship between intelligence and inspection time concluding that the test shows excellent and high correlations with IQ. Bates and Eysenck (1993) have used both tests of RT and of IT to correlate with IQ, showing that they measure distinct aspects of IQ.

Jensen (1982b), Brand and Deary (1982) and many others attempt to explain the observed relationships between the different RT paradigms and IQ in terms of speed of mental operations. As Jensen points out, the conscious brain acts as a one-channel or *limited capacity* information processing system which can deal simultaneously only with a very limited amount of information. This limited capacity also restricts the number of operations that can be performed simultaneously on the information that enters the system from external stimuli, or from retrieval of information stored in short-term or long-term memory. "Speediness of mental operations is advantageous in that more operations per unit of time can be executed without overloading the system" (p. 122). Jensen goes on to point out that there is *rapid decay* of stimulus traces and information, so that there is an advantage to speediness of any operation that must be

performed on the information while it is still available. Third, he goes on to say, the individual resorts to *rehearsal and storage* of the information into intermediate or long-term memory, which has relatively unlimited capacity, in order to compensate for limited capacity and rapid decay of incoming information. However, the process of storing information in LTM itself takes time and therefore uses up channel capacity, so that there is a trade-off between the storage and the processing of incoming information. The more complex the information and the operations required on it, the more time is required, and consequently the greater the advantage of speediness in all elemental processes is involved. "Loss of information due to overload interference and decay of traces that were inadequately encoded or rehearsed for storage or retrieval from LTM results in "breakdown" and failure to grasp all the essential relationships among the elements of a complex problem needed for its solution. Speediness of information processing, therefore, should be increasingly related to success in dealing with cognitive tasks to the extent that the information loss strains the individual's limited channel capacity" (p. 122).

Jensen goes into this theory in much greater detail than can be done here; I will return to the theory and suggest an alternative hypothesis when we have dealt with the work on evoked potential and other measures of psychophysiological processes basic to the measurement of intelligence. Here let us merely note that both the theory developed by Jensen, and the facts as briefly outlined here are quite incompatible with a Binet-type theory. If differences in intelligence are produced essentially by environmental factors related to education, experience, cultural influences, and so on, or to the development of strategies for coping with problems of a cognitive kind, then there should be no correlation at all, or at best a minimal one, between RT and IQ. The fact that correlations are in fact consistent and high is one of those anomalies which prove fatal to an established paradigm, and suggest the need for, not just revision, but a complete revolution. As we shall see, this need will become much more urgent when we consider the psychophysiological literature in relation to IQ measurement, but even the RT data by themselves would be sufficient to indicate that such models of intelligence as those put forward by Sternberg (1982) and his fellow authors are unacceptable because they cannot accommodate the RT data. Sternberg does not argue the case in his Handbook; he simply presents his "metatheory of intelligence" without mention of the contrary evidence presented by RT and A.E.P. data. For other recent Binet-type theories, see Snow et al. (1976) and Chiang and Atkinson (1976).

EEG POTENTIALS AS MEASURES OF INTELLIGENCE

It might have been hoped that, if we take Galton's theory of a physiological measure of intelligence seriously, that the EEG would present us with good

evidence in this respect. However, a careful survey by Lindsley (1961) showed quite early on that there is no clear-cut relation in normal adults. (See also Andreassi, 1980; Gale & Edwards, 1983.) In early childhood, where slower gamma and delta waves give way to fast prevalent alpha, and in brain-injured persons (Ostow, 1954), some relationships are found. Among younger children and mental defectives, or deteriorated adults with a mental age of between 6 and 10, correlations can be found ranging from .3 to .6 between mental age and more rapid, developed alpha rhythm. Reviews by Ellington (1966) and Vogel and Broverman (1964, 1966) are contradictory, the former denying on the whole any relationship in normal adults between EEG and intelligence, the latter putting forward evidence that in normal children mental age and alpha frequency are correlated, when age and sex are held constant. Callaway's (1975) summary is probably in closest agreement with the fact when he says that the EEG-intelligence correlations reported thus far are promising rather than nonexistent, but that the repetitive wavelike aspects of the ongoing EEG are not the best windows on the mind. Such repetitive waves in the EEG tell us more about what the mind is *not* doing than about what the mind is doing. Wavelike activities often signal a failure to operate, and if the brain is usually busy on a variety of jobs, the EEG is a jumble of signals appropriate to the jumble of underlying processes, and hence not likely to reflect dispositional qualities like *g*.

More recent work on the spontaneous EEG has been surveyed by Deary and Caryl (1993). The most important studies are those of Giannitrapani (1985) and Gasser et al. (1983). Clearly quite significant correlations with IQ of the order of 0.5 can be obtained, but the whole field lacks proper theoretical underpinning. Apparently activity in *particular* narrow frequency bands may be important here, for example, 13Hz rather than the higher amplitude, neighboring dominant alpha-frequency of 11Hz. Different topographic sites also play an important role. Coherence may also be very relevant (Thatcher et al., 1983), with *lower* coherence, and hence more differentiated brain activity, being associated with high intelligence.

Instead of relying on ordinary EEGs as measures of intelligence, there has in recent years been a spurt of studies using event-related potentials, such as the contingent negative variation (CNV), and more importantly, the averaged evoked potential (AEP). There is little evidence that CNV is correlated with intelligence; it is more of the nature of a measure of attention or arousal, and will not be dealt with in detail here. In passing, however, it may be said that work on the CNV is suggestive, because its failure to correlate with intelligence may indicate that attention and arousal are not active components in those correlations with intelligence that have been found on the AEP. As will be shown, these correlations are probably highest in situations which do not call forth any high degree of attention or arousal. This is important because it contradicts a widely-held view that possibly the relationship between AEP and intelligence may be mediated by factors of arousal and attention.

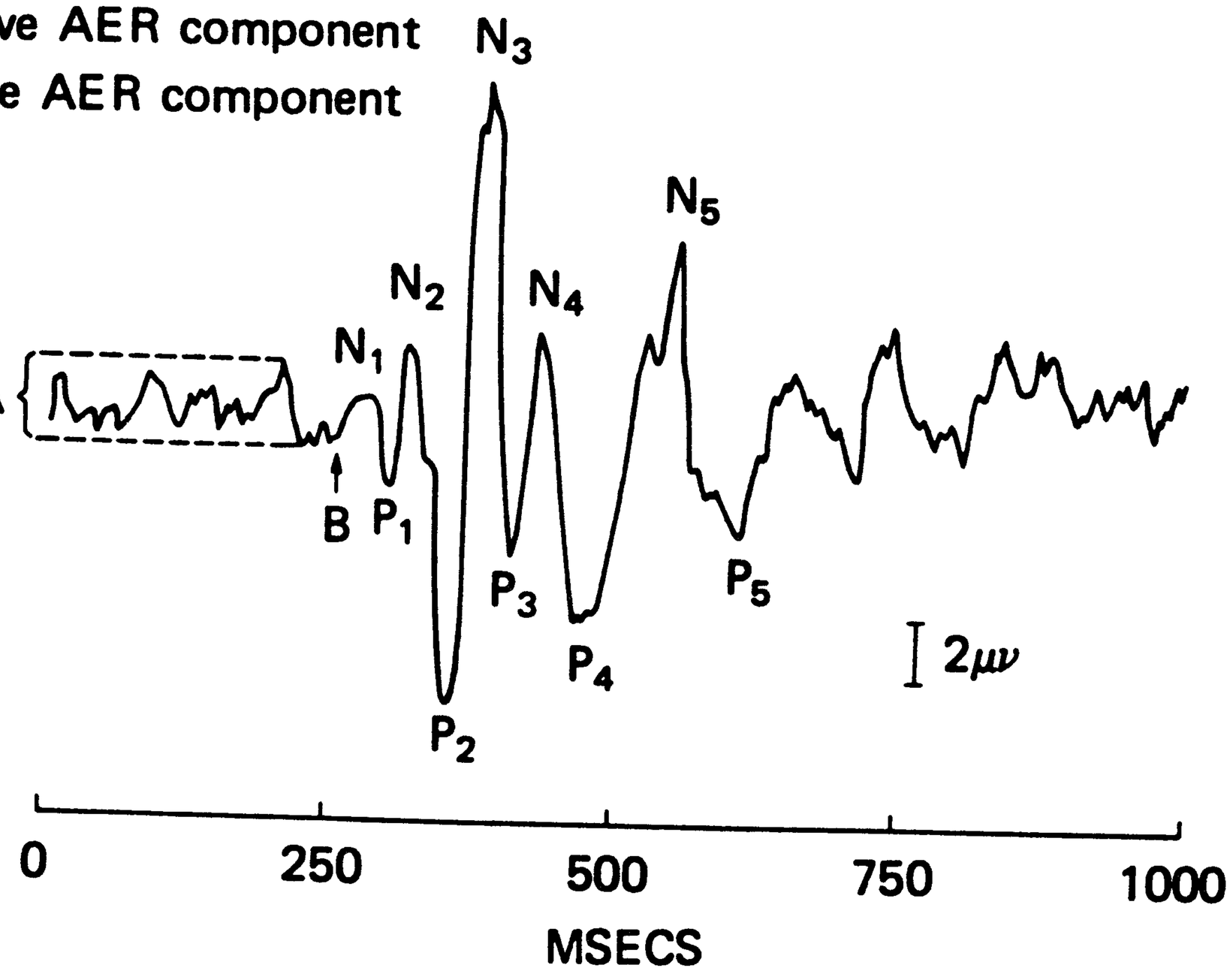


Figure 5. Averaged Evoked Potential Responses

AEPs are essentially a series of negative and positive potentials which are produced by some kind of sensory stimulus, usually visual or auditory. Figure 5 shows in very diagrammatic form such a series of waves following the onset of a stimulus (A), which is preceded by a series of ordinary EEG resting level waveforms. The AEP returns to normal after something like 500 to 1000 milliseconds, but for the purpose of measurement of intelligence it is the first 250 milliseconds or so which are particularly important. Because of the poor signal to noise ratio of the AEP, which is only partly corrected by making the stimulus phase-dependent on the resting EEG, a number of time locked evocations have to be averaged in order to give a recognizable and measurable wave (Shagass, 1972; Regan, 1972; Basar, 1980). As John (1973) has pointed out, such averaging may in fact obscure event-related potentials—so-called noise may not be noise but only poorly understood signals. Nevertheless, averaging has been widely used, and has become indispensable for the study of AEPs.

The wave-form of the AEP differs from person to person, depending on the nature of the stimulus, electrode location, intensity of the stimulus, and the state of the subject. The earlier components of the AEP are usually referred to as far-field components, they reflect bipoles set up by receptors and peripheral relay stations (Jewett et al., 1970). Far-field potentials are very small and have very short latencies.

Following the far-field components are the so-called primary cortical evoked responses; these are called near-field, since small changes in electrode position may have large effects, unlike the far-field components. They are also characterised by larger amplitude and later occurrence (from 20 to 100 milliseconds). A detailed discussion of AEP characteristics in relation to intelligence measurement is given by Eysenck and Barrett (1985), and Barrett and Eysenck (1982), and we will not here repeat the details, but will concentrate rather on substantive matters which emerge from the early work, before turning to more recent, theory-oriented and successful applications.

These early data, in particular, show a great divergency of results, due in large part to the fact that different investigators used different electrode placements, bipolar versus unipolar recordings, different types of stimuli, different intensities of stimuli, different interstimulus intervals, different instructions, different types of waveform analyses, and other parameter differences which may be related to attentional and similar psychological factors; the resulting scores, usually different from one investigator to another, were then correlated with different types of IQ test, some measures of fluid, some measures of crystallised ability and some more closely resembling measures of educational learning. Under such conditions, confusion is clearly likely to reign supreme, and we can at best expect to disentangle to some modest extent the most promising types of parameter arrangements likely to give positive results in connection with IQ measurement.

In Binet's approach to the testing of intelligence, the concept of mental age played a vital part, that is, the notion that intelligence increases with age, so that any mental test on which scores increase with age may be a likely candidate for inclusion in a battery of IQ tests. A similar approach, as we have seen, can be made in the field of reaction time, where also RTs and variability of RT decrease with increasing age. Callaway (1975) gives a good survey of a large body of evidence to show that certain aspects of the AEP change from early to late childhood, and might thus be candidates for inclusion in a battery of mental age, and possibly of IQ measures. In particular, with increasing age there is a reduction in the variability of AEP, and a reduction also in *latency*. This agrees surprisingly well with the findings on RTs, where also there is a reduction in *variability*, and a reduction in reaction time. These might therefore present plausible avenues to explore, and as we shall see positive results have been obtained. However, it should be added that in actual fact age-related measures are not necessarily correlated more highly with IQ than are non-age correlated measures, so that these relationships with age should be regarded as suggestive, and not as conclusive.

Chalke and Ertl (1965) reported the first correlations between AEP measures and IQ. They found short visual AEP latencies correlated with high IQ in a group of 48 rather heterogeneous subjects, including 4 mentally retarded subjects. This early work was taken up in several other papers by Ertl and

Schafer (1969), and Ertl (1969, 1971, 1973). Many frequency-domain measures were tried in an attempt to improve correlations, such as Fourier analysis, complex demodulation, analogue spectral analysis, and so fourth, but the time-domain measure has remained the measure of choice. All this work may perhaps be summarised by saying that on the average correlations between latency and standard IQ test scores had been found in the neighborhood of -.30 in a fairly consistent fashion.

Work done by Hendrickson (1972) in my own Department gave some support for Ertl's conclusions; her unpublished data are summarised by Eysenck (1973). Using 93 adults, randomly sampled, she administered the AH4 test of intelligence, which gives a verbal, spatial and a total score; she also determined latencies and amplitudes of evoked potentials in response to sounds of three different intensities. Both latency (negatively) and amplitude (positively) were found correlated with intelligence, more so with verbal than with spatial intelligence, and most of all with total intelligence score. The average size of the correlations ranged from .30 to .50 for latency and from .30 to .45 for amplitude, when we are considering verbal ability, and from .10 to .25 when we are considering spatial ability. When it is borne in mind that latency and amplitude are essentially uncorrelated, and that we can therefore add variance estimates to predict IQ, then we may say that quite roughly such a combined score of latency and amplitude would correlate with IQ between .50 and .60; if this value were corrected for attenuation, it would give a correlation between .60 and .70.

Latency and; to a lesser extent, amplitude have been the main parameters of the AEP to be investigated, but others have also played a part. *Variability* in the evoked potential may be even more important. As evoked potentials are averaged, the AEP of course contains no information about the variability of the individual measures on which it is based. This variability is inevitably connected with amplitude, greater variability reducing amplitude (Brazier, 1964). Variability may also account for increase in latency, so that amplitude, latency and variability are not independent (Callaway, 1975). It is possible that variability, rather than being something to be got rid of, may be a more fundamental factor than other variables, and as we shall see, this view has been successfully followed by Hendrickson in a study to be mentioned in the next section.

EP variability *decreases* in children with age from 6 to 16, which suggests a negative correlation between variability and IQ. This hypothesis has been shown by Callaway (1975) to give correct predictions, as in his own work with children; low EP variability is usually associated with superior performance. In later work with large samples of navy recruits, Callaway also found similar results for adults, but correlations tended to be rather lower, centering around .20. The predictability of IQ in terms of EP variability may be *underestimated* by the product moment correlation. In plotting his data he found a nearly

empty quadrant where poor EP stability and good IQ performance would be located. He suggests that it is as though the opportunity to learn and other factors play an increasingly important role in limiting performance once a neural substrate has attained sufficient stability. In other words, the neural processes producing stable evoked potentials are *necessary* but not *sufficient* for superior mental performance.

The last major approach to the study of AEP-IQ correlations to be mentioned here has been by way of AEP *asymmetry*. Asymmetrical AEPs can be produced by somatosensory stimulation, resulting in a contralateral parietal AEP that has more complex early components than does the ipsilateral parietal AEP. Similarly, one-sided stimulation of other modalities produces asymmetrical AEPs. Cognitive operations can also be used in this context, propositional (left hemisphere) cognitive operations tending to suppress the EEG from the left hemisphere and to reduce AEPs to irrelevant stimuli, with appositional (right hemisphere) tasks suppressing right hemisphere EEGs and task irrelevant AEPs.

Greater asymmetry has been found in bright children by Rhodes et al. (1969), Lairy et al. (1969), with two unpublished Ph.D. studies quoted by Callaway (1975) failing to find significant differences. For adults an early study by Giannitrapani (1969), and a more recent one by Callaway (1975), have noted positive results. The latter only found correlations under non-attending conditions, and he again observed the phenomenon on the "empty quadrant." However, in later studies he failed to find a correlation between asymmetry and IQ.

While there is considerable variety and disagreement in the results reported in these early studies, certain features stand out. *High IQ seems to be correlated significantly with short latencies, large amplitudes, and low variability*; these are the major aspects of evoked potential performance to have given positive results, and as already pointed out, these three features of the EP performance are not independent. Variability might for many reasons be regarded as the most fundamental; a high variability precludes great amplitude and short latencies. There is no suggestion in this early work on the reasons for some individuals to show greater ability, nor is there any attempt to link up this factor of variability correlations with lower IQ, with the similar observation in the reaction time literature also linking high variability with low IQ.

It is typical of these early studies in the field that there is little theoretical underpinning for the empirical results. Most authors seem to have chosen parameter values almost at random, and in disregard of theoretical problems, and this may in part be responsible for the low values of the correlations found, which typically account for no more than 10 percent of variance. Perhaps workers in the field might have benefited from Kurt Lewin's advice, stating that there is *nothing as practical as a good theory*. This lack of a theory should be taken together with the fact that no two investigators seem to have used

identical or even similar conditions. In view of their differences in the positioning of the electrodes, the use of unipolar or bipolar electrode placements, the intensity of the stimuli, their modality, the arrangement of trials and inter-trial intervals, instructions given to subjects, internal states of motivation, and many other aspects of the situation, it seems a miracle that even such small amount of agreement as has been found did in fact emerge from such varied studies. Deary and Caryl (1993) give an excellent account of all this early work.

THE HENDRICKSON PARADIGM

The work of A.E. and D.E. Hendrickson has been based on a theoretical model of memory, information processing and intelligence which goes into considerable biochemical and neurological details involving neuron functioning and events at the synapse. This model (A.E. Hendrickson, 1982; D.E. & A.E. Hendrickson, 1980) is too complex to be described in any detail in this chapter; a more extended discussion of it is given by Eysenck and Barrett (1985). Another reason for not presenting the model in detail is that whether the many hypotheses incorporated in it are correct or not does not essentially affect the validity of the empirical findings concerning the relationship between AEP and intelligence. Inevitably the model is highly speculative, and only future research will be able to tell to what extent it resembles reality, if at all. Our main concern here is with deductions made from this model regarding those measures of AEP performance which might be expected to correlate with intelligence, and we will not attempt here to formulate criticisms or defences of the model as such.

Consider the actual AEP recordings of three children, one bright, one average, one dull (see Figure 6). The shorter latencies of the bright child, and the longer latencies of the dull child, will be obvious. Amplitudes are not so obviously different. But what is apparent to the naked eye is the greater *complexity* of the waveform of the bright child, and the *simplicity* of the waveform of the dull child. It is this feature of the AEP that is basic to the Hendricksons' approach. Their argument, essentially, is that *variability* produces simplicity in the waveform, and is in turn produced by *errors in neural processing* occurring in the central nervous system. To obtain complexity in the (averaged) waveform, successive AEP response curves have to be very similar, trough on trough, and peak on peak. Errors produce differences in the response curves, and hence only the major sinusoidal curves remain; all finer, smaller features are obliterated.

Variability is one measure suggested by Hendrickson's theory; the other relates to the *complexity* of the waveform, assessed by measuring the contour perimeter of the AEP wave shape. The more intelligent the individual,

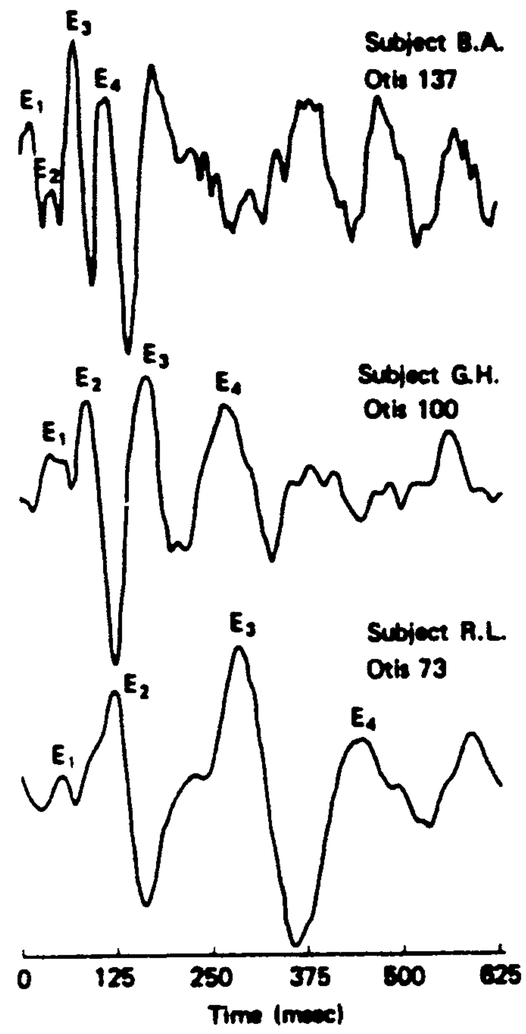


Figure 6. Averaged Evoked Potential Responses of a Bright, Average and Dull Child

the longer the contour or, in the terms used previously, the less "smooth" will be the appearance of the AEP wave shape. This complexity measure is sometimes called the "string" measure because in the early feasibility studies undertaken by the Hendricksons the wave shape was projected onto a sheet of paper attached to a block of wood, pins were driven into the wood at all points where the curve changed direction, and then a string was threaded around these pins following the actual form of the wave. Finally the string was taken off the pins, and its length measured; this gives a rough and ready measure of the "complexity" of the wave form. Needless to say, computer methods of measurement have since replaced this primitive but nevertheless very useful method of measurement. Figure 7 illustrates, in terms of typical AEP curves of six bright and six dull children, the differences in "complexity" which are clearly apparent even to the naked eye.

The theory would lead us to expect that variability should correlate negatively, complexity (the string measure) positively with IQ, and that these two measures should correlate negatively but quite highly with each other. There are two studies bearing on these points. The first (Blinkhorn & Hendrickson, 1982) correlated the complexity measure with performance on Raven's Advanced Progressive Matrices, and a variety of verbal ability tests. Auditory stimuli were used for 33 student subjects. Various correlations between the string measure and IQ were calculated, yielding a midrange correlation of approximately 0.45. Verbal test scores did not show significant correlations. Because of the restricted range of the scores on the Progressive Matrices Test, Blinkhorn and Hendrickson corrected this value assuming a full range of IQ, thus boosting the correlation to a maximum of 0.84.

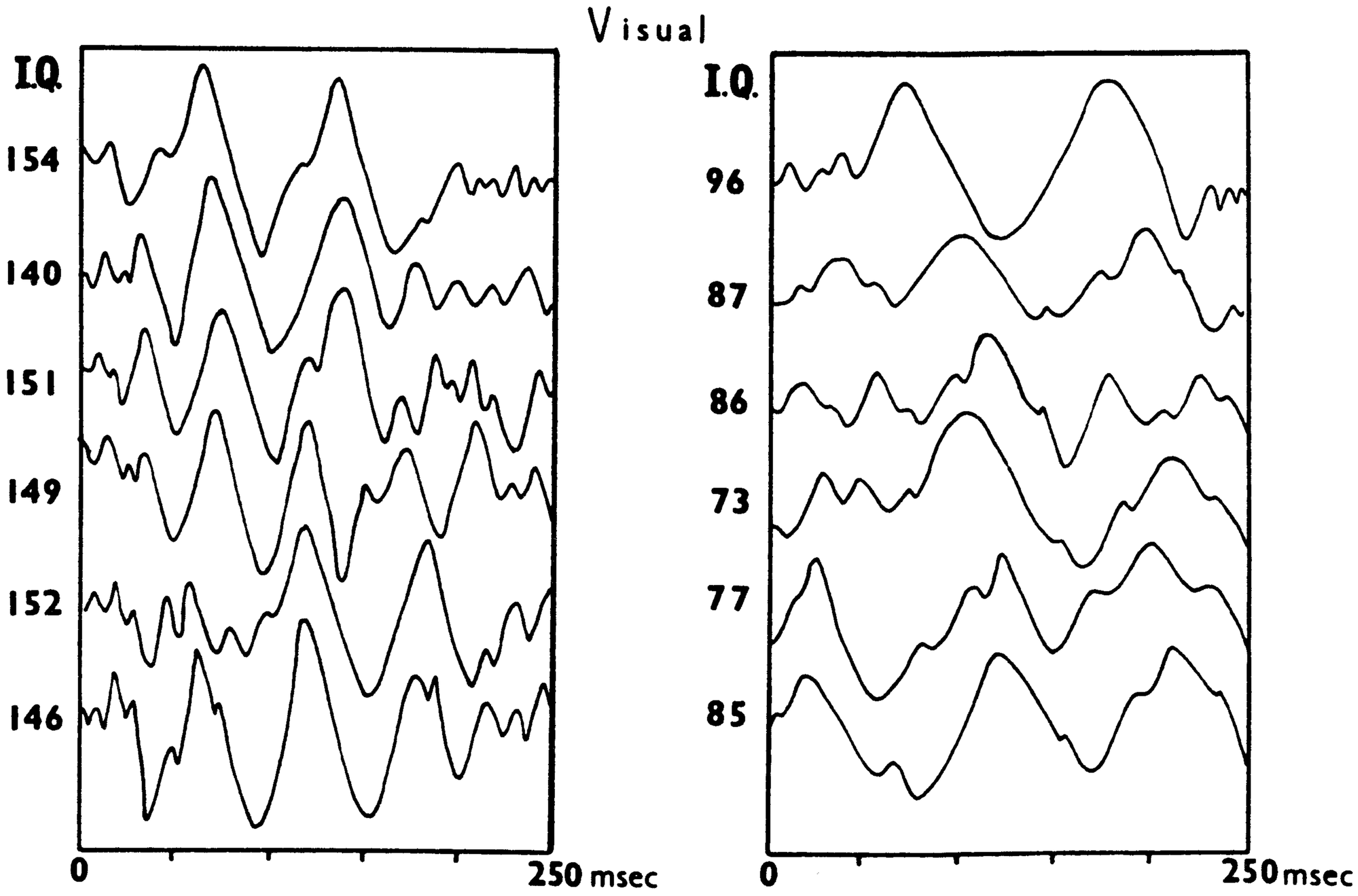


Figure 7. Averaged Evoked Potential Responses of 6 Bright and 6 Dull Children

This value is reasonably similar to one obtained by the Hendricksons (1980) in an analysis of some published data of Ertl's, which showed a correlation of 0.77 between WISC IQ scores and the string score from the evoked potentials, as measured by the Hendricksons. However, only extreme IQ scores were reported in the Ertl study, thus probably exaggerating the degree of correlation present.

The second study (D.E. Hendrickson, 1982) is much more extensive and impressive. Using a reasonably random sample of 219 school children, she administered the WAIS (Wechsler, 1955) to assess IQ, scores being given for separate subscales, a performance total, a verbal total, and an overall IQ total score. In addition, to the complexity and variance measures defined above, Hendrickson used a *composite* measure composed of the variance score minus the string score. Table 2 shows the results of this experiment.

As will be seen from Table 2, the correlation between the IQ and string, variance and composite AEP measures are 0.72, -0.72, and -0.83 respectively. The correlations between the WAIS performance total and the string, variance and composite measures are -0.53, 0.53 and -0.60 respectively. The correlations between the WAIS verbal total and the string, variance and composite measures are 0.68, -0.69 and -0.78 respectively. Thus, unlike the Blinkhorn and Hendrickson study, it appears that the AEP measures correlate higher with verbal ability than performance related abilities.

Table 2. Relationship Between the EEG Measures and the WAIS Subtests

WAIS test	*Variance*	*String*	*Variance minus string*	*Full WAIS IQ (current study)*	*Full WAIS IQ (published data)*
Information	-.64	.55	-.68	.80	.84
Comprehension	-.50	.53	-.59	.74	.72
Arithmetic	-.57	.56	-.65	.79	.70
Similarities	-.69	.54	-.71	.84	.80
Digit span	-.54	.49	-.59	.71	.61
Vocabulary	-.57	.62	-.68	.79	.83
Verb. total	-.69	.68	-.78	.95	.96
Digit symbol	-.28	.32	-.35	.45	.68
Picture comprehension	-.47	.52	-.57	.67	.74
Blocks	-.50	.45	-.54	.70	.72
Picture arrangement	-.36	.45	-.46	.54	.68
Object assembly	-.32	.45	-.44	.55	.65
Performance total	-.53	.53	-.60	.69	.93
WAIS total	-.72	.72	-.83	1.00	1.00

A factor analysis was carried out, using the 11 WAIS scales and the composite AEP score (Eysenck & Barrett, 1985). Only one general factor was extracted, as this would represent in a direct form the g factor common to all the tests. On this factor the AEP measure had a loading of 0.77 (in a previous analysis, carried out by Eysenck, 1982b, data were used from a slightly different group of subjects, including as well as the schoolchildren, some high IQ adults. Hence the results of these two analyses are not identical). In another analysis correlations for the WAIS subtests uncorrected for attenuation, and also corrected for attenuation are given; the correction was performed in order to answer the argument that possibly the higher correlations of some tests might be due to their higher reliabilities. Of particular interest is the Spearman Rho Correlation between the factor loadings and the composite AEP measure of the 11 subscales of the WAIS. If we accept the psychometric assumption that this factor represents in an optimal fashion the general intelligence factor as measured by the widely varying subtests of the WAIS, and if we assume that the composite AEP measure is a relatively pure measure of this central general factor, then we would expect the set of 11 factor loadings, whether corrected or uncorrected, to correlate with the composite AEP measures, whether corrected or uncorrected. In other words, tests having high factor loadings (i.e., being good measures of general intelligence) should show high correlations with a composite AEP measure, while tests having low factor loadings should have low correlations with the composite AEP measure. There is a notable degree

of correspondence between the two sets of figures, the correlation between the uncorrected figures being 0.95, and that between the corrected figures being 0.93. These figures are remarkably high, considering that the range of factor loadings is not very large (due to the fact that all tests obviously were only incorporated in the Wechsler because they were reasonably good measures of *g*.) We may conclude, therefore, that *whatever is measured by the subtests of the WAIS is measured also by the composite AEP measure.*

Many people have carried out replications of these studies, with varying degrees of success (Deary & Caryl, 1993, give an excellent review). The appearance of the occasional "failure to replicate" has encouraged studies to look at possible *causes* for such failure, and Barrett and Eysenck (1993) have found a possible cause in the N100-P180 component of the AEP waveform, that is, the amplitude of the positive deviation of the waveform at 180 msec. as compared with the negative deviation of 100 msec. This finding is too recent to have been subjected to the necessary replication and critique, but it suggests important ways of interpreting the original findings, and the outcomes of the many replications.

Other paradigms have also been proposed and investigated, in particular the Schaffer paradigm. In recent years, several studies have demonstrated a cognitive modulation of EEG activity, with particular reference to the amplitude of AEPs. Hillyard et al. (1973) and Picton and Hillyard (1974) have demonstrated the influence of selective attention; Squires et al. (1976) and Schafer and Marcus (1973) have demonstrated the influence of expectancy, and Schafer (1978) and Israel et al. (1980) have performed the same service for information processing workload. It appears that unexpected or "attended" stimuli produce AEPs of larger overall amplitude than those generated using stimuli whose nature and timing is known to the individual. Schafer has extended the scope of this empirical phenomenon, suggesting that *individual differences in the modulation of amplitude will relate to individual differences in intelligence.* He has labelled this modulation of amplitude "cognitive neural adaptability," and the physiological basis mediating this relationship is hypothesized to be neural energy as defined by the number of neurons firing in response to the stimulus. A functionally efficient brain will use fewer neurons to process a foreknown stimulus, whereas for a novel, unexpected stimulus, the brain will commit large numbers of neurons. Given the relationship between individual neuron firing patterns and observed cortical AEPs, the commitment of neural energy will be observed as amplitude differences between AEPs elicited from various stimulus presentation conditions.

It follows from this general line of reasoning that individuals with high neural adaptability, characterized by AEPs with much smaller than average amplitude to expected stimuli and much larger than average amplitude to unexpected stimuli, should show high intelligence on IQ tests, whereas conversely, for individuals with low neural adaptability, the degree of such AEP amplitude

modulation should be diminished, with a corresponding low intelligence test score. Studies by Schafer and Marcus (1973) and Schafer (1982) have given strong support to their theory, which may possibly be linked with the Hendrickson one by assuming that in people with cerebral integrity repeated stimuli are more rapidly recognised as repetitive. For yet other paradigms, Deary and Caryl (1993) or Barrett and Eysenck (1993) may be consulted.

CONCLUSION

Modern work strongly supports the Galton paradigm, while admitting some merit in the Binet approach. Essentially it is now believed that the integrity of the central nervous system is fundamental to the proper functioning of the cortex; that errors in the transmission of information lower cortical efficiency, as expressed in the IQ, and that such errors also reduce mental speed. They can be measured as a function of the variability of the AEP, which also finds expression in the string length. These theoretical advances may link the DNA, basic to cognitive ability, with behavior expressed in terms of the IQ. This nested group of theories constitutes the "new look" in intelligence (Eysenck, 1982a, 1987).

REFERENCES

Andreasen, N.C., Flaum, M., Swayze, V., O'Leary, D.S., Alliger, P., Cohen, G., Ehsferoll, J., & Yuk, W. (1993). Intelligence and brain structure in normal individuals. *American Journal of Psychiatry*, *150*, 130-134.

Andreassi, J.L. (1980). *Psychophysiology: Human behavior and physiological response.* New York: Oxford University Press.

Barnes, B. (1982). *T.S. Kuhn and social science.* London: McMillan.

Barrett, P., & Eysenck, H.J. (1982). Brain electrical potentials and intelligence. In Gale, A. & M. Eysenck (Eds.), *Handbook of individual differences: Biological perspectives* (pp. 255-285). London: Wiley.

Barrett, P., & Eysenck, H.J. (1993). The relationship between evoked potential component amplitude, latency, contour length, variability, zero-crossings, and psychometric intelligence. *Personality & Individual Differences*, *16*, 3-32.

Basar, E. (1980). *EEG-Brain dynamics.* New York: Elsevier.

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

Baumeister, A.A., & Kellas, G. (1968). Distribution of reaction times of retardates and normals. *American Journal of Mental Deficience*, *72*, 5, 715-718.

Blinkhorn, S.F., & Hendrickson, D.E. (1982). Averaged evoked responses and psychometric intelligence. *Nature*, *295*, 596-597.

Brand, C.P., & Deary, I.J. (1982). Intelligence and inspection time. In Eysenck, H.J. (Ed.), *A model for intelligence.* New York: Springer.

Brazier, M.A.B. (1964). Evoked response recorded from the depths of the human brain. *Annals of the New York Academy of Science*, *112*, 33-59.

Brody, N. (1992). *Intelligence*. San Diego: Academic Press.
Burt, C. (1909). Experimental tests of general intelligence. *British Journal of Psychology*, *3*, 94-177.
Callaway, E. (1975). *Brain electrical potentials and individual psychological differences*. London: Grune & Stratton.
Carroll, J. (1993). *Human cognitive abilities*. Cambridge: Cambridge University Press.
Cattell, R.B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*, 1-22.
Cattell, R.B. (1971). *Abilities: Their structure, growth and action*. Boston: Houghton Mifflin.
Chalke, F., & Ertl, J. (1965).Evoked potentials and intelligence. *Life Sciences*, *4*, 1319-1322.
Chiang, A., & Atkinson, R.C. (1976). Individual differences and interrelationships among a select set of cognitive skills. *Memory and Cognition*, *4*, 661-672.
Deary, I., & Caryl, P. (1993). Intelligence, EEG, and evoked potentials. In Vernon, P.A. (Ed.), *Biological approaches to the study of human intelligence* (pp. 259-316). Norwood, NJ: Ablex.
Ellington, R.J. (1966). Relationship between EEG and test intelligence: A commentary. *Psychological Bulletin*, *65*, 91-98.
Ertl, J.P. (1969). Neural efficiency and human intelligence. Final Report. *U.S. Office of Education Project*, No. 9-0105.
Ertl, J. (1971). Fourier analysis of evoked potentials and human intelligence. *Nature, 230*, 525-526.
Ertl, J. (1973). IQ evoked responses and Fourier analysis. *Nature*, *241*, 209-210.
Ertl, J., & Schafer, E. (1969). Brain response correlates of psychometric intelligence. *Nature*, *223*, 421-422.
Eysenck, H.J. (1939). Primary mental abilities. *British Journal of Educational Psychology*, *9*, 270-275.
Eysenck, H.J. (1967). Intelligence assessment: A theoretical and experimental approach. *British Journal of Educational Psychology*, *37*, 81-98.
Eysenck, H.J. (Ed.). (1973). *The measurement of intelligence*. Lancaster: MTP.
Eysenck, H.J. (1979). *The structure and measurement of intelligence*. New York: Springer.
Eysenck, H.J. (1982a). *A model for intelligence*. New York: Springer.
Eysenck, H.J. (1982b). The psychophysiology of intelligence. In Spielberger, C.D. & Butcher, J.M. (Eds.), *Advances in personality assessment* (Vol. 1). Hillsdale, NJ: Lawrence Erlbaum.
Eysenck, H.J. (1983). Revolution dans la theories et la mesure de l'intelligence. *La Review Canadienne de Psycho-Education*, *12*, 3-17.
Eysenck, H.J. (1987). Speed of information processing, reaction time, and the theory of intelligence. In Vernon, P.A. (Ed.), *Speed of information processing and intelligence* (pp. 21-67). Norwood, NJ: Ablex.
Eysenck, H.J. (1988). The concept of "intelligence": useful or useless? *Intelligence*, *12*, 1-16.
Eysenck, H.J. (1992). Intelligence: The one and the many. In D.K. Detterman (Ed.), *Current topics in human intelligence: Is mind modular or unitary?* (Vol. II, pp. 83-116). Norwood, NJ: Ablex.
Eysenck, H.J. (1993). The biological basis of intelligence. In Vernon, P.A. (Ed.), *Biological approaches to the study of human intelligence*, (pp. 1-32). Norwood, NJ: Ablex.
Eysenck H.J., & Barrett, P. (1985). Psychophysiology and the measurement of intelligence. In Reynolds, C. & Willson, V. (Eds.), *Methodological and statistical advances in the study of individual differences*. New York: Plenum.
Frearson, W., & Eysenck, H.J. (1986). Intelligence, reaction time (RT) and a new "odd-man-out" RT paradigm. *Personality & Individual Differences*, *7*, 807-817.
Gale, A., & Edwards, J.A. (1983). Individual differences. In Coles, M., Donchin, E. & Porges, S. (Eds.), *Psychophysiology: Systems, processes and applications*. New York: Guilford Press.
Galton, F. (1883). *Inquiries into human faculty and its development*. London: Macmillan.
Galton, F. (1879). *Hereditary genius*. London: Julian Friedman.

Gasser, T., Lucadou-Muller, I., Verleger, R., & Bacher, P. (1983). Correlating EEG and IQ.: A new look at an old problem using computerised EEG parameters. *Electroencephalography and Clinical Neurophysiology*, *55*, 493-509.

Giannitrapani, D. (1969). EEG average frequencies and intelligence. *Electroencephalography and Clinical Neurophysiology*, *27*, 480-486.

Giannitrapani, D. (1985). *The electrophysiology of intellectual function*. Basle: S. Harper.

Guilford, J.P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Guilford, J.P., & Hoepfner, R. (1971). *The analysis of intelligence*. New York: McGraw-Hill.

Hendrickson, D.E. (1972). *An examination of individual differences in the cortical evoked response*. London: Unpublished Ph.D. Thesis, University of London.

Hendrickson, D.E. (1982). The biological basis of intelligence. Part II: Measurement. In Eysenck, H.J. (Ed.), *A model for intelligence*. New York: Springer Verlag.

Hendrickson, D.E., & Hendrickson, A.E. (1980). The biological basis of individual differences in intelligence. *Personality and Individual Differences 1*, 3-33.

Hick, W. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*, 11-26.

Hillyard, S.A., Hink, R.F., Schwent, V.L., & Picton, T.W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*, 177-180.

Hunt, E. (1976). Varieties of cognitive power. In Renick, L.B. (Ed.), *The nature of intelligence*. Hillsdale, NJ: Erlbaum.

Israel, J.B., Wickens, C.D., Chesney, G.L., & Donchin, E. (1980). The event related brain potential as an index of display monitoring workload. *Human Factors*, *22*, 211-224.

Jensen, A.R. (1982a). The chronometry of intelligence. In Sternberg, R.J. (Ed.), *Advances in the psychology of human intelligence*. London: Lawrence Erlbaum.

Jensen, A.R. (1982b). Reaction time and psychometric g. In Eysenck, H.J. (Ed.), *A model for intelligence*. New York: Springer.

Jensen, A.R., & Johnson, F.W. (1994). Race and sex differences in head size and IQ. *Intelligence*, *18*, 309-333.

Jensen, A.R., & Munro, E. (1974). Reaction time, movement time, and intelligence. *Intelligence*, *3*, 121-126.

Jensen, A.R., Schafer, E.W.P., & Crinella, F.M. (1981). Reaction time, evoked brain potentials and psychometric g in the severely retarded. *Intelligence*, *5*, 179-197.

Jensen, A.R., & Sinha, S.N. (1993). Physical correlates and human intelligence. In Vernon, P.A. (Ed.), *Biological approaches to research in human intelligence*. Norwood, NJ: Ablex.

Jewett, D.L., Romano, M.N., & Williston, J.S. (1970). Human auditory evoked potentials: Possible brain stem components detected on the scalp. *Science*, *107*, 1517-1518.

John, E.R. (1973). Brain evoked potentials: acquisition and analysis. In Thompson, R.F. & Patterson, M.M. (Eds.), *Bioelectric recording techniques*. New York: Academic Press.

Keating, D.P., & Bobbitt, B. (1978). Individual and developmental differences in cognitive processing components of mental ability. *Child Development*, *49*, 155-169.

Kuhn, T.S. (1957). *The Copernican revolution*. Cambridge, MA: Harvard University Press.

Kuhn, T.S. (1970). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Kuhn, T.S. (1974). Second thoughts as paradigms. In Suppe, F. (Ed.), *The structure of scientific theories*. Illinois: Illinois University Press.

Lairy, G.C., Remond, A., Rieger, H., & Lesevre, N. (1969). The alpha average. III. Clinical application in children. *Electroencephalography and Clinical Neurophysiology*, *26*, 453-467.

Lehrl, S. Gallwitz, A., & Blaham, L. (1980). *Kurztest Für allgemeing intelligence kai*. Munich: Manual Vlass.

Lally, M., & Nettlebeck, T. (1977). Intelligence, reaction time, and inspection time. *American Journal of Mental Deficiency*, *82*, 273-281.

Lemmon, V. (1927). The relation of reaction time to measure of intelligence, memory and learning. *Archives of Psychology*, *15*, 5-38.

Lindsley, D.B. (1961). The reticular motivation system and perceptual integration. In Sheer, D.E. (Ed.), *Electrical stimulation of the brain.* Austin: University of Texas Press.

Masterman, M. (1970). The nature of the paradigm. In Lakatos, I. & Musgrave, A. (Eds.), *Criticism and the growth of knowledge.* Cambridge: Cambridge University Press.

McCauley, C., Dugas, J., Kellas, G., & Devellis, R.F. (1976). Effects of serial rehearsal training on memory search. *Journal of Educational Psychology*, *68*, 474-481.

Ostow, M. (1954). Psychodynamic disturbances in patients with temporal lobe disorder. *Journal of Man / Sinai Hospital*, *20*, 293-308.

Peake, H., & Boring, E.G. (1926). The factor of speed in intelligence. *Journal of Experimental Psychology*, *9*, 71-94.

Picton, T.W., & Hillyard, S.A. (1974). Human auditory evoked potentials: effects of attention. *Electroencephalography and Clinical Neurophysiology*, *36*, 191-199.

Plomin, R., & McCleary, G. (Eds). (1993). *Nature, nurture and psychology.* Washington, DC: American Psychological Association.

Posner, M.I. (1969). Abstraction and the process of recognition. In Bower, G.H. & Spence, J.T. (Eds.), *The psychology of learning and motivation* (Vol. 3, pp. 43-100). New York: Academic Press.

Regan, D. (1972). *Evoked potentials in psychology, sensory physiology and clinical medicine.* New York: Wiley-Interscience.

Rhodes, L., Dustman, R., & Beck, E. (1969). The visual evoked response: a comparison of bright and dull children. *Electroencephalography and Clinical Neurophysiology*, *27*, 364-372.

Roth, E. (1964). Die geschwindigkeit der verarbeitung von information und ihr zusammenhang mit intelligenz. *Zeitschrift fur Experimentelle und Angewandte Psychologie*, *11*, 612-622.

Schafer, E.W.P. (1982). Neural adaptability: a biological determinant of behavioural intelligence. *International Journal of Neuroscience*, *17*, 183-191.

Schafer, E.W.P., & Marcus, M.M. (1973). Self stimulation alters human sensory brain responses. *Science*, *181*, 1175-1177.

Shagass, C. (1972). *Evoked brain potentials in psychiatry.* New York: Plenum Press.

Shugard, D., & Horn, J. (1972). Evoked critical potential and measurements of human abilities. *Journal of Comparative and Physiological Psychology, 78*, 59-68.

Snow, R.E., Marshalek, B., & Lohman, D.F. (1976). Correlation of selected cognitive abilities and cognitive processing parameters: an exploratory study. Technical Report 3. *Aptitude Research Project.* Palo Alto: Stanford University School of Education.

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

Squires, K.C., Wickens, C., Squires, N.K., & Donchin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event related potential. *Science*, *192*, 1142-1146.

Sternberg, R.J. (Ed.). (1982). *Handbook of human intelligence.* London: Cambridge University Press.

Sternberg, S. (1966). High speed scanning in human memory. *Science*, *1553*, 652-654.

Sully, J. (1876). Physiological psychology in Germany. *Mind*, *1*, 20-43.

Thatcher, D., McAlaster, R., Lester, M., Horst, R., & Cantor, D. (1983). Hemispheric EEG asymmetries related to cognitive functioning in children. In *Cognitive processing in the right hemisphere* (pp. 125-146). New York: Academic Press.

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

Thurstone, L.L., & Thurstone, T.G. (1941). *Factorial studies of intelligence.* University of Chicago Press.

Vernon, P.E. (1979). *Intelligence: Heredity and environment.* San Francisco: Freeman.

Vernon, P.E. (1981). Reaction time and intelligence in the mentally regarded. *Intelligence*, *5*, 345-355.

Vogel, W., & Broverman, D.M. (1964). Relationship between EEG and test intelligence: A critical review. *Psychological Bulletin*, *62*, 132-144.

Vogel, W., & Broverman, D.M. (1966). A reply to "Relationship between EEG and test intelligence: a commentary." *Psychological Bulletin*, *65*, 99-109.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. New York: Psychological Corporation.

Wissler, C. (1901). The correlation of mental and physical tests. *Psychological Review Monographs*, *3*.

Wolf, T.H. (1973). *Alfred Binet*. Chicago: University of Chicago Press.

THE MENTAL SPEED APPROACH TO THE ASSESSMENT OF INTELLIGENCE

A. C. Neubauer

HISTORY OF THE MENTAL SPEED APPROACH

The notion of human intelligence as an inherited feature of the central nervous system which can be measured by means of simple sensory or sensorimotor tasks dates back to Galton (1883). In the late nineteenth century, he devised a battery of such tasks measuring reaction time, sensory acuity and physical energy, and tested several thousand subjects in his "anthropometric laboratory." The results, however, were disappointing: Simple reaction time (RT) and sensory discrimination abilities did not show any relationship to independent signs of accomplishment or intelligence. Nevertheless, Galton's idea stimulated further research efforts: Wissler (1901) used 21 different "mental tests" devised by his supervisor J. McKeen Cattell, but could not find any substantial correlation with the college grades of 300 students. Similar results were obtained by Sharp (1898-1899). As Deary (1994) points out, these two studies were frequently cited as grounds for dismissing the idea of a relationship between mental speed and intelligence. More importantly, they were also

Advances in Cognition and Educational Practice, Volume 4, pages 149-173.

ISBN: 0-7623-0105-8

highly influential on research practice, in that they effected an almost complete abandonment of this line of research until Erwin Roth, a German psychologist in Erlangen, revived the idea in 1964. He began with Hick's (1952) discovery of a linear relationship between the amount of information in a visual RT task and the RT of a subject: In the simple RT there is no choice and, therefore, no information involved (0 bits of information), so deciding between two alternatives (visual stimuli) requires one binary decision (1 bit); when four alternatives are presented, two binary decisions (bits) are necessary and so on. Hick demonstrated that the RT bears a logarithmic relationship to the number of alternatives, which becomes linear when using the logarithm (to the base 2) of the number of stimulus alternatives.

On the basis of Hick's discovery, Roth (1964) tested 58 subjects with an RT task known today as the Hick paradigm, that is, he employed a visual RT task under eight conditions of varying informational load (from simple RT to eight choice RT corresponding to 0-3 bits of information). He hypothesized the slope of the regression of RTs on bits of information to be an indicator of *speed of information processing* (SIP), which should correlate negatively with psychometric intelligence (more intelligent subjects should need less time to process one bit of information). The -.39 correlation between this slope measure and the Intelligenz-Struktur-Test (I-S-T) not only confirmed his hypothesis but also served as a frequently quoted starting point for modern studies relating SIP in so-called elementary cognitive tasks (ECTs) to psychometric intelligence.

EMPIRICAL EVIDENCE

It took another 15 years for Roth's findings to stimulate a further attempt to relate RTs in an ECT to psychometric intelligence. Jensen and Munro (1979) were only partially successful in replicating Roth's observation: However, the correlation of the Hick RT slope with psychometric intelligence (-.30) was in the expected direction but failed to reach significance ($p < .06$). Other parameters like mean RTs or the intraindividual variability (*SD*) of RT produced largely higher correlations (up to -.49) with Raven's Standard Progressive Matrices (SPM, Raven, 1960). This study resulted in major research efforts to relate RTs in the Hick paradigm (and other ECTs) to psychometric intelligence.

The Hick Paradigm

The majority of the studies relating parameters of the Hick paradigm to psychometric intelligence employed three to four conditions of stimulus information: simple RT (0 bit), two-choice RT (1 bit), four-choice RT (2 bits)

Table 1. Means of Hick Parameters

	Bits of Information					
	0	*1*	*2*	*3*	*Intercept*	*Slope*
Median RT	335	373	401	439	337	34
SDRT	46	47	58	84	39	5

Source: From Jensen (1987a).

and eight-choice RT (3 bits). In most of the reports the following parameters were correlated with psychometric intelligence: Median RTs and intraindividual variabilities (*SD*) of RTs for each bit-condition as well as intercept and slope of the regression of median RTs on bits of information. An overview of all studies until 1987 relating the Hick paradigm to psychometric intelligence is given by Jensen (1987a). From 33 independent samples comprising a total of 2317 subjects, he calculated means of the Hick parameters (see Table 1) as well as their mean correlations with psychometric intelligence (see Table 2).

Table 1 shows that the median RTs correspond quite well with Hick's law; the mean correlation between bits of information and median RT is .995. The intraindividual variability SDRT also increases with bits ($r = .943$), but a better correspondence ($r = .991$) can be obtained when regressing the SDs onto the number of response alternatives instead of bits of information.

When these parameters are correlated with psychometric intelligence, the coefficients rarely exceed -.40; in the majority of the studies correlations between -.10 and -.30 were observed. Table 2 summarizes the mean correlations calculated by Jensen (1987a). Regarding the modest size of the coefficients, it must, however, be remarked that most of the studies used relatively homogeneous samples of an above-average intellectual level, which probably has led to an underestimation of the true relationship between Hick parameters and intelligence.

Table 2 shows expected as well as paradoxical findings. The increase of RT-IQ correlations with increasing informational load agrees well with early findings by Lemmon (1927) and Roth (1964) and—according to Jensen—can be explained by referring to Spearman's (1927) concept of general intelligence (*g*): The more complex an ability test the better it is suited to assess *g*.

Table 2. Mean Correlations of Hick Parameters with Psychometric Intelligence

	Bits of Information					
	0	*1*	*2*	*3*	*Intercept*	*Slope*
Median RT	-.18	-.19	-.22	-.23	-.12	-.12
SDRT	-.26	-.21	-.28	-.25	-	-

Source: From Jensen (1987a).

The fact that RT-intelligence correlations increase with informational load (and therefore complexity) can be viewed as evidence that ECTs capture primarily aspects of *g*. An unexpected finding in Table 2 is the low correlation between the slope of RT and intelligence (-.12). On the basis of information processing theory this parameter should be substantially (if not directly) related to psychometric intelligence. Jensen attributes this negative finding to the weak stability of this parameter (mean test-retest reliability = .39, Jensen, 1987a).

The highest correlations with psychometric intelligence, however, can be observed for the parameters of intraindividual variability of RT (see Table 2). Jensen explains this finding by referring to the typical skewness of intraindividual RT distributions: As there is a physiological limit for the shortest possible RTs, individual differences in mean RT can be caused only by the individual distribution of longer RTs. Baumeister and Kellas (1968) demonstrated that retarded vs. normal subjects differ only slightly in their shortest RTs but largely in their slowest RTs. On the basis of this evidence, Jensen assumes that individual RT differences are not caused primarily by some kind of SIP but by the "moment-to-moment probability of occurrence of certain processes" (1987a, p. 134). As a physiological basis for this "moment-to-moment probability," Jensen (1982) suggests an "oscillation rate" of neurons or groups of neurons (see section 4).

The Inspection Time (IT) Paradigm

The measurement of "Inspection Time" (IT; Vickers, Nettelbeck, & Willson, 1972) was also frequently used in the mental speed approach to human intelligence. In this visual discrimination task, a simple visual stimulus consisting of two vertical lines of different length is presented to the subject for very short time durations by means of a tachistoscope (see Figure 1a), with a subsequent masking of the stimulus (Figure 1b). The subject's task is to decide which one of the two vertical lines is longer. By assessing the probabilities of correct responses to varying stimulus exposure times (usually ranging from about 10 to 200 msec), the minimum required time for a near-perfect discrimination (e.g., 97.5%) can be measured; this is the individual IT. Usually the individual IT is estimated (if necessary by means of interpolation) from the function relating stimulus exposure times to the probability of correct responses. Vickers et al. (1972) observed ITs ranging from 55 to 169 msec ($M = 106$, $SD = 30$).

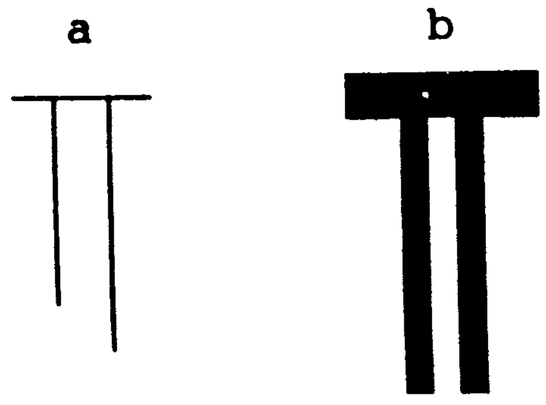


Figure 1. The 2-lines Discrimination Task for Measuring Inspection Time: (a) Target figure; (b) Masking figure

Studies relating ITs to psychometric intelligence are reviewed by Kranzler and Jensen (1989). Most of the studies reviewed followed the basic rationale of Vickers et al. (1972), but some of them also used different sensory modalities, masking procedures, response criteria and so on. In most of these studies negative IT-intelligence correlations of moderate size were observed. Table 3 shows the results of Kranzler and Jensen's meta-analysis (1989). They distinguished studies relating IT to measures of general, performance and verbal intelligence (IQ), as well as studies involving adults' vs. children's samples. Moreover, coefficients corrected for sampling error, attenuation and range restriction are given. The substance of Table 3 is a negative correlation between ITs and intelligence of moderate size (around -.3, uncorrected). The correlations are lower for verbal and higher for performance measures of IQ. After averaging the corrected correlations in adults, Kranzler and Jensen (1989) conclude the "true" IT-IQ correlation to be -.54. Although this might be a rather optimistic estimate (due to the fact that three corrections were applied to the original coefficients), a moderate but substantial negative relationship between IT and psychometric intelligence can be taken for granted.

Sternberg's Short-Term Memory Scanning

The study of the relationship between Hick RTs and ITs on the one hand and psychometric intelligence on the other was an attempt to isolate some kind of SIP (or mental speed) as an elementary basis of human intelligence. Both the Hick and the IT-paradigm measure primarily the speed of perceiving and encoding (visual) stimuli; in addition the Hick paradigm also measures speed of response choice. Notwithstanding the assumed utility of this approach, Jensen conjectured back in 1982 that—beyond the basic mental speed— the velocity of certain other elementary cognitive processes (ECPs) might also constitute a fundamental source of individual differences in human intelligence. Among the ECPs he proposed was the speed of information-retrieval from short-term and long-term memory.

A well-known task to assess the speed of retrieval from short-term memory is Saul Sternberg's "short-term memory scanning" paradigm (1969). In this paradigm the subject is sequentially presented with a random sequence of digits (usually varying in length between 1 and 5 digits), which have to be kept in short-term memory (STM). After a warning stimulus, a target item (digit) is displayed and the subject has to indicate as quickly as possible if the digit was in the previously shown memory set. The typical finding from such an experiment is a linear increase of RT with the number of elements in STM. This linear relationship allows us to calculate the regression of RTs on the number of items in STM. The slope of this regression shows the time needed for the STM-retrieval of a single element. The intercept of the regression gives the time for the duration of encoding and motoric response processes.

Table 3. Results of Kranzler and Jensen's (1989) Meta-analysis of IT-intelligence Studies

			Corr. With IT	
Mental Test Sample	*N*	*K*	*Uncorrected*	*Corrected*
General IQ				
Adults	633	17	-.31	-.56
Children	242	8	-.37	-.59
Performance IQ				
Adults	88	4	-.49	-.74
Children	230	7	-.32	-.47a
Verbal IQ				
Adults	218	7	-.26	-.38
Children	248	8	-.29	-.41

Note: [a] No range correction

Early attempts to relate performance in the Sternberg paradigm to psychometric intelligence mainly compared groups of different intellectual ability. Dugas and Kellas (1974) and Harris and Fleer (1974) observed significantly steeper slopes and higher intercepts in mental retardates than in subjects of average intelligence. They concluded that retardates are characterized by a slower access to STM as well as by a slowing of stimulus encoding and/or motoric reaction. Group comparisons in the "normal range" of intelligence confirmed the slope differences (McCauley, Dugas, Kelley, & DeVellis, 1976; Keating & Bobbitt, 1978) but reported only tendential differences in the intercept.

Studies using a correlational approach are reviewed in Table 4. For this survey all currently available studies reporting zero-order correlations between parameters from the Sternberg paradigm and psychometric intelligence are included.[1] The average correlations in Table 4 show that it is mainly the mean (or median) RTs, the SDs of RT and the intercept which correlate moderately with psychometric intelligence (between -.27 and -.35). Contrary to expectations, however, are the low correlations of the slope measure with IQ. This could be taken as a strong argument against a prominent role of speed of STM-access for human intelligence, but for the moment I will not comment further on that idea. A thorough explanation will be advanced later. Generally, there is sound evidence for a substantial negative relationship between Sternberg-RTs and RT-variabilities on the one hand and psychometric intelligence on the other.

Posner's "Letter Matching" Paradigm

Jensen (1982) also hypothesized the speed of long-term memory (LTM) access to be an important predictor of individual differences in human intelligence.

Table 4. Survey of Studies Reporting Zero-order Correlations of the Sternberg Paradigm with Psychometric Intelligence (IQ)

		Correlations With IQ			
Study	*Sample/ IQ-test*	*Mean RT*	*SDRT*	*Slope*	*Intercept*
Lunneborg, 1977	64 school-ch./ EAS			-.14[1]	-.32[1]
Jenkinson, 1983	60 children/ SPM	-.36[2]		.06	-.33
McGue et al., 1984	105 twins/ WAIS			-.26[3]	-.27[3]
Cohn et al., 1985	60 gifted children 70 non-gifted ch. SPM	-.42 -.14			
Vernon et al., 1985	81 students/ MAB			-.05[3]	
Vernon and Kantor, 1986	113 students/ MAB	-.40[4]			
Jensen, 1987b	48 students/ APM	-.30	-.37	-.06	-.32
Kranzler, 1991	101 students/ APM & MAB	-.23	-.34		
Deary et al., 1992	100 hypoglycemia-patients WAIS	-.33			
Miller and Vernon, 1992	170 students/ MAB	-.11[5]			
mean correlation:		-.29	-.36	-.09	-.31
mean correlation	(N-weighted):	-.27	-.35	-.11	-.30

Notes: For the computation of mean correlations Fisher's Z-transformation was used
[1] average from correlations with 8 cognitive ability tests
[2] average from correlations with 3 set sizes
[3] average from the correlations of the slope for positive and for negative trials with IQ
[4] average from two groups
[5] average from correlations with 10 cognitive ability tests
Abbreviations of psychometric IQ-tests:
EAS: Employee Aptitude Survey, SPM: Raven's Standard Progressive Matrices
APM: Raven's Advanced Progressive Matrices, WAIS: Wechsler Adult's Intelligence Scale,
MAB: Multidimensional Aptitude Battery

In mental speed research most researchers employed Posner and Mitchell's (1967) "letter matching" task to assess this ability: In each trial the subject is simultaneously shown two letters, which are either physically the same (e.g., "AA"), physically different but semantically same ("Aa") or physically and semantically different ("Ab"). In the first trial the subject judges the Physical Identity (PI-test), in the second trial he judges the semantic or Name Identity (NI-test). While the PI-test only demands visual discrimination, the NI-test additionally requires access to highly overlearned material stored in LTM (the letters of the alphabet). The difference between the mean RTs in the NI- and PI-tests (NI-PI) should, therefore, indicate the time needed for LTM retrieval (Hunt, 1980).

Early attempts to relate parameters of the Posner paradigm to psychometric intelligence compared groups of varying intellectual (mostly verbal) ability. In a series of studies Hunt and coworkers found smaller NI-PI differences in groups of high verbal ability, indicating a shorter LTM access in verbally proficient subjects (cf. Hunt, 1980; but also Goldberg, Schwartz, & Stewart, 1977; Keating & Bobbitt, 1978).

Studies using a correlational approach (and reporting zero-order correlations between parameters from the Posner paradigm and psychometric intelligence) are surveyed in Table 5.[2] The mean correlations are lowest for the mean RT in the PI-test, which is not surprising as this is a less complex condition compared to the NI-test, which shows a higher mean correlation. As expected, the NI-PI difference measure also correlates negatively with intelligence, but as in the early studies (which compared groups) the generalizability of this finding is questionable since only two of the five studies reporting NI-PI by IQ correlations used measures of general intelligence; of the remaining three, two used only verbal ability tests and one a measure of crystallized intelligence. The implications of this restriction for the mental speed theory will be dealt with later.

To conclude this section, it should be mentioned that other ECTs were also employed in mental speed research, but due to the small number of such studies, they will not be reviewed here.

IQ-Prediction from Multiple Elementary Cognitive Tasks

While the early mental speed studies mainly employed only single ECTs, more recent studies attempted to predict general intelligence from a battery of ECTs. As already outlined in 2.3, the basic idea behind this procedure is that complex problem solving (like in intelligence tests) not only requires abasic mental speed (as measured by the Hick paradigm) or a speed of

Table 5. Survey of Studies Reporting Zero-order Correlations of the Posner Paradigm With Psychometric Intelligence (IQ)

		Correlations With IQ		
Study	*Sample/ IQ-test*	*PI-RT*	*NI-RT*	*NI-PI*
Warren, 1978 (after Hunt, 1980)	Children (*N*=?)/ WISC-Verbal IQ			-.34[1]
Jackson and McClelland, 1979	52 students/ School & College Aptitude Test	-.22[2]	-.36[2]	
Hunt et al., 1981	69 students/ Nelson-Denny Test			-.25[1]
Lansman et al., 1982	91 students/ diverse			-.35[3]
McGue et al., 1984	105 twins/ WAIS	-.27	-.44	-.12
Cohn et al., 1985	60 gifted children 70 non-gifted ch./ SPM	-.25 -.25	-.52 -.25	
Vernon et al., 1985	81 students/ MAB			-.39[4]
Vernon and Kantor, 1986	113 students/ MAB	-.30	-.50	
Kranzler, 1991	101 students/ APM & MAB	-.18	-.21	
Miller and Vernon, 1992	170 students/ MAB	-.17[5]	-.22[5]	
Vernon and Weese, 1993	152 students/ MAB		-.27	
mean correlation: mean correlation	(*N*-weighted):	-.23 -.23	-.35 -.33	-.29 -.27

Notes: For the computation of mean correlations Fisher's Z- transformation was used
[1] Correlation with verbal IQ
[2] average from correlations with verbal und numerical cognitive ability
[3] correlation with "crystallized intelligence"
[4] average from correlations with verbal and performance IQ and from two samples
[5] average from correlations with 10 cognitive ability tests
Abbreviations of psychometric IQ tests:
WISC: Wechsler Intelligence Scale for Children
(for the remaining abbreviations see Table 4)

sensory discrimination (assessed by the IT task) but also a variety of other elementary cognitive processes (like speed of STM and LTM retrieval). Therefore, a combination of several ECTs (via multiple regression) should result in a better prediction of psychometric intelligence. A first attempt in this direction can be ascribed to Keating and Bobbitt (1978): In a sample of 60 children a combination of choice RT, Sternberg and Posner paradigms resulted in a multiple correlation of .72 with Raven's matrices test. This figure, however, could be an overestimation because the sample was very heterogeneous with respect to age. But even within the three age groups tested (9, 13 and 17 years) multiple R's around .58 were observable.

Similar studies involving adults gave less impressive results. For a combination of simple and choice RTs, IT, Sternberg and Posner paradigms, Vernon (1983) found only a multiple *R* of .46 in 100 undergraduate students. Probably due to the multicollinearity of the ECTs, this relationship scarcely exceeded the correlations of single ECTs with intelligence (.38 to .45). The same problem of multicollinearity of ECTs seems to account for Jensen's similar findings (1987b): The combination of the Hick paradigm, a visual search task and the Sternberg paradigm yielded a multiple *R* of .50; the largest zero-order correlation of a single ECT variable with intelligence was -.46.

Integration of RT-Intelligence Evidence: The Complexity Hypothesis

The empirical evidence presented so far has shown that the size of the mean correlations of ECTs or their parameters with intelligence seems to depend on the complexity of the ECT: More complex ECTs (i.e., those with longer mean RTs) correlate more highly with intelligence than less complex ones. This phenomenon has also received empirical support in a study by Vernon and Jensen (1984) in which they employed eight different ECTs and reported an almost perfect relationship between the mean latency of an ECT and its negative correlation with psychometric intelligence ($r = .96$ after reflection of the RT-IQ correlations), that is, the more complex an ECT the higher its correlation with psychometric intelligence. Since Vernon and Jensen's first report, the complexity phenomenon has been frequently replicated (e.g., Cohn, Carlson, & Jensen, 1985; Vernon & Weese, 1993). An explanation of the complexity hypothesis advanced by Larson, Merritt and Williams (1988) refers to Detterman's (1986) definition of general intelligence, according to which *g* can be defined as the universe of cognitive operations. Correlations with *g* reflect the extent to which the universe is sampled. This has been confirmed in the psychometric domain, where Marshalek, Lohman and Snow (1983) demonstrated that more complex intelligence tests, which require more cognitive operations, display higher loadings on *g* than less complex tests. The higher correlations of more complex ECTs with psychometric intelligence seem

to parallel findings in the psychometric domain; moreover, the higher correlations suggest that ECTs primarily measure *g*. From this it can further be concluded that ECTs should show the highest correlation with psychometric tests known to be highly *g*-loaded. Evidence related to this question shall be presented in the next section.

SPEARMAN'S "GENERAL INTELLIGENCE" AND MENTAL SPEED

In fact, most proponents of the mental speed approach assumed that the SIP is the elemental basis of individual differences in general intelligence or *g* (in Spearman's sense, 1927). Therefore, most of the studies relating ECTs to psychometric intelligence used intelligence tests which are known to load highly on *g*, such as Raven's Standard and Advanced Progressive Matrices (1960, 1958) and the Wechsler scales for adults (WAIS = "Wechsler Adult Intelligence Scale") and for children (WISC = "Wechsler Intelligence Scale for Children"). However, the use of such intelligence tests does not allow a stringent test of the hypothesis that ECTs primarily measure aspects of *g*. This hypothesis can only be tested by employing a battery of intelligence tests and demonstrating that more highly *g*-loaded intelligence tests also show a higher correlation with ECT parameters than intelligence tests with a low *g*-loading.

Such an attempt was first reported by Hemmelgarn and Kehle (1984). They tested 95 children with the WISC, submitted the intercorrelations of the 12 WISC subtests to a factor analysis and determined the subtests' *g*-loadings by extracting the first unrotated factor (the loadings on this factor are usually considered good indices of the *g*-saturation of an intelligence test, Jensen, 1992). They could show a close correspondence between the *g*-loadings and the RT-intelligence correlations, that is, the higher the *g*-loading of a WISC subtest, the larger its (negative) correlation with parameters from the Hick paradigm.

This first evidence that ECTs predominantly assess aspects of general intelligence and not more specific intellectual abilities has been replicated by several authors (Larson et al., 1988; Smith & Stanley, 1987; Vernon, 1989). A critic of this position, Ceci (1990) contrasts this "singularity of mind" view with the "specificity of mind" hypothesis, which—briefly—states that correlations between microlevel efficiency (ECTs) and macrolevel measures (intelligence tests) are due only to their sharing a common knowledge base. Following this view, a verbal ECT should correlate only with verbal intelligence, and the same should be true for numerical or visual contents. Empirical tests of this alternative view, which correlated verbal, numerical and visual intelligence tests with mental speed tests involving these three kinds of stimuli, did not, however, support Ceci's view (Levine, Preddy, & Thorndike, 1987; Lindley, Bathurst, Smith, & Wilson, 1993; Neubauer & Bucik, 1996).

The correlations between ECTs and intelligence tests assumed to have a common knowledge base were only marginally different from ECT-intelligence correlations involving different knowledge bases.

THEORY OF THE MENTAL SPEED APPROACH

In the previous sections it has been demonstrated that there is sound empirical evidence for a moderate but consistent negative relationship between the speed of elementary cognitive processes and psychometric intelligence. How can this relationship be explained theoretically? For such a theoretical explanation most proponents of the mental speed approach refer to Jensen's "oscillator model" and/or to Vernon's "neural efficiency hypothesis." Only the first will be outlined here.[3]

In the introduction to his model, Jensen (1982) refers to several "classical" concepts of cognitive psychology. The first basic assumption is that the human brain is a one-channel information processing system with a limited capacity. Second, information required for problem solving may not only come from external sources but also from the internal short term or long term memory (STM or LTM). The limited capacity restricts the number of operations that can be performed on information from such external or internal sources. Therefore, a high SIP has the advantage that more cognitive operations can be performed per unit of time. This assumption alone, however, cannot explain why a high SIP facilitates problem solving even in untimed intelligence tests. To explain this Jensen refers to the *rapid decay* of stimulus information: As stimulus information is not available for an unlimited amount of time, quick information processing has the advantage that more cognitive operations can be performed on the stimuli while they are still available. Finally, the large amount of information that has to be handled in complex tasks (items of an intelligence test) makes it necessary to store parts of the information in the LTM, which has relatively unlimited capacity. The process of storing information in LTM takes time and, therefore, "competes" with the processing of incoming information.

The more complex the information and the operations required by the task, the more advantageous is a high speed of execution of elementary cognitive processes (ECPs), because it minimizes the danger of losing information, which itself could effect a "breakdown" of information processing. Such a breakdown can result in "a failure to grasp all the essential relationships among the elements of a complex problem needed for its solution" (Jensen, 1982, p. 122). Therefore, the speed of ECPs (like processing incoming information, information retrieval from and storage in STM and LTM) should determine the individual point of breakdown in performance of a complex task. According to Jensen, the speed of these ECPs can be best measured by means of the ECTs described in section 2, because these are so simple that a breakdown can be ruled out.

As a physiological basis for individual differences in the speed of elemental processes, Jensen proposes two properties of the central nervous system: (1) the number of neural elements activated by a stimulus and (2) the rate of excitatory-refractory phases in activated neural elements (= oscillation rate). According to Jensen the second property in particular is a central determinant of RTs; for example for the Hick paradigm, Jensen assumes that the number of dichotomous decisions (bits) must be isomorphically represented in the neural networks of the cerebral cortex, that is, each decision corresponds to one node of excitation, which will fire when a critical level of stimulation is reached. As the level of excitation at each node oscillates, it must be refractory (inactive) half of the time, and during this phase excitation cannot be transmitted to the next node, causing delays. RTs in the Hick paradigm should therefore depend mainly on two factors: (1) the number of nodes through which the excitation must be conducted and (2) the mean oscillation rate of the transmitting nodes. The first factor explains the linear increase of RT with bits of information. The second is the source of individual differences in RTs: The higher the oscillation rate, the faster the transmittal of excitation through the nodes and therefore the shorter the RT. Moreover, the concept of oscillation rate can explain individual differences in RT variability (SDRT): In a series of RT trials a higher oscillation rate leads to smaller variations in transmittal of excitation and therefore in RT.

In his model Jensen explicitly refers to a biological substrate to explain the correlations between performance in ECTs and psychometric intelligence. As this is also true for other mental speed-theories (the "neural efficiency" model by Vernon, 1993, and the biochemical approach by A.E. & D.E. Hendrickson, 1982), these are sometimes labeled "bottom-up" approaches as opposed to "top-down" explanations, which will be presented in the next section.

TOP-DOWN EXPLANATIONS OF THE MENTAL SPEED APPROACH

Critics of the mental speed approach emphasize the role of high-level cognitive processes, like attention, motivation, learning ability and others. Some of these factors might influence not only RTs in ECTs but might also correlate with psychometric intelligence like learning ability (Guthke, 1993) or attentional resources (Westhoff & Kluck, 1983). Therefore, empirically observed RT-intelligence correlations might be mediated by such high-level cognitive processes. As these alternative viewpoints emphasize the role of higher level cognitive processes, they can be labeled "top-down" explanations. In the following the most detailed top-down approaches by Longstreth and R.J. Sternberg and other criticism will be summarized along with empirical evidence.

Longstreth's Criticism of the Hick Paradigm

Longstreth's (1984) criticism mainly focused on the classical Hick apparatus and procedure used by most researchers:

1. *Order effects*: In the classical procedure the levels of information (set sizes) are always presented in their order of magnitude (from 0 to 3 bits of information). The amount of information, therefore, is compounded with the order of trials, which could result in an accumulation of practice effects from smaller to larger set sizes and therefore in an underestimation of the RT to larger set sizes and of the slope of RT. As one could assume "a positive relationship between IQ and rate of learning" (Welford, 1986, p. 194), this would artificially raise the slope-IQ correlation.

2. *Visual attention effects*: In the classical Hick procedure the magnitude of the visual field controlled by the subject varies with the amount of information. A retinal displacement of the stimulus effects longer RTs and as the magnitude of retinal displacement is correlated with bits of information, an overestimation of the slope could be the result.

3. *Response practice effects*: There is a compounding between the amount of information and the number of certain responses/movements required from the subject. As the set sizes were always presented in ascending order, the subject has practised the responses required during the smaller set sizes more when receiving the larger set sizes.

Empirical studies designed to test Longstreth's hypotheses, however, did *not* provide evidence that these sources of artifact had am important impact on RTs and/or on RT-intelligence correlations (Kranzler, Whang, & Jensen, 1988; Larson & Saccuzzo, 1986; Neubauer, 1990a; Widaman & Carlson, 1989). The author (Neubauer, 1991) even found higher than usual RT-intelligence correlations for a modified Hick paradigm, which avoids the sources of artifact described by Longstreth.

R. J. Sternberg's Criticism

On the basis of the experiential subtheory of his triarchic theory of intelligence, R.J. Sternberg (1985; Marr & Sternberg, 1987) assumes that neither extremely novel tasks nor completely automatized tasks constitute good measures of intelligence. For an extremely novel task, the subject first has to develop the appropriate mental representation of the problem (early metacomponential processes), so that he is initially unable to apply his full range of abilities to the problem. On the other hand, for a completely automatized task, performance depends primarily upon the speed and accuracy of a preselected sequence of performance components, with a negligible contribution from metacomponents. In RT-IQ research a small number of practice trials are usually given which, according to Sternberg, ensure that when

the experiment starts the task is neither extremely novel nor completely automatized for the subject, that is, he is in a state where task-relevant metacomponents are involved. This could explain the observed negative RT-IQ correlations, but at the same time predicts a decrease in correlations with increasing practice (and therefore automatization) on the RT task. A similar prediction is made by Ackerman and Schneider (1985) on the basis of Schneider and Shiffrin's framework of controlled vs. automatic processing (1977) as performance on relatively simple RT tasks can be assumed to become automatized with practice.

Neubauer and Freudenthaler (1994) tested this hypothesis of declining RT-IQ correlations with increasing practice by letting 60 undergraduates perform more than nine hours (or 2688 trials) on a well-known ECT, the so-called Sentence-Verification-Test (SVT; Clark & Chase, 1972). In spite of a considerable reduction of RTs after 9 hours (about one third), the (negative) RT-intelligence correlations (with Raven's APM) remained fairly constant (see Figure 2). This finding clearly contradicts top-down explanations of the RT-intelligence relationship on the basis of metacomponential processes (Sternberg, 1985) or controlled vs. automatic processing (Ackerman & Schneider, 1985). Instead, from our findings we infer strong support for the biologically based bottom-up explanations of the mental speed-theory of intelligence.

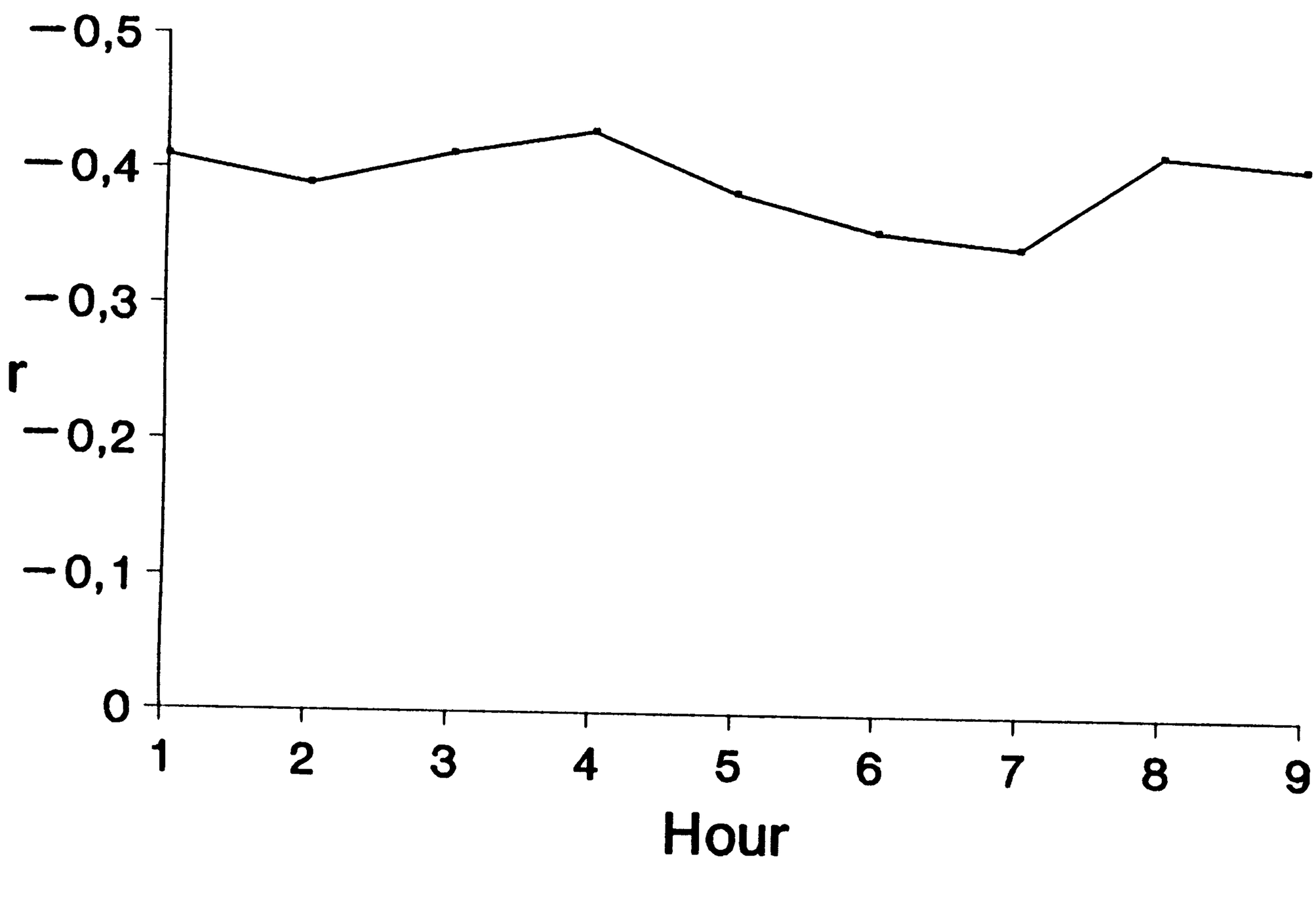


Figure 2. Time-Course of Correlations Betweeen SVT-RTs and Psychometric Intelligence (APM)

The Role of Attention

Some authors (Carlson & Widaman, 1987; Detterman, 1987; Longstreth, 1986) assume that empirically observed RT-intelligence correlations might only be explained by individual differences in attention or concentration ability. They assume that low-IQ subjects are not paying as much attention in RT experiments as high-IQ subjects, perhaps because "they do not recognize the importance of paying attention in order to minimize reaction time" (Longstreth, 1986, p. 187) and this would have the effect of inflating the correlations between RT and IQ. In fact Carlson and Widaman (1987) reported evidence for an association between high attention and fast and consistent RTs but—with respect to their finding of near zero correlations between RTs and intelligence—their study seems nonrepresentative considering the large number of studies demonstrating such a relationship.

A more recent attempt to elucidate the role of attention in RT-intelligence relations was reported by the author (Neubauer, Bauer & Hoeller, 1992). Two samples of school children ($N = 81$ and 125) were tested with Raven's SPM (1960), the d2 attention test (Brickenkamp, 1978) and with Neubauer's (1991) modified Hick paradigm. Due to large individual differences in the set for speed versus accuracy, an information measure (average information rate in bits/sec) was calculated which combines speed (mean RT) and accuracy (number of errors). This information measure correlated .29 and .36 (in Studies I and II, respectively) with the SPM and .47 and .29 with the d2 (all four correlations significant with $p < .01$). Therefore, SIP in the Hick paradigm was associated with both intelligence and attention. The hypothesis that the RT-IQ relationship is mediated by individual differences in attention was tested by computing a first-order partial correlation between the information measure (from the Hick paradigm) and the SPM controlling for the d2: In both samples the partial correlations were only somewhat lower than the zero-order correlations and still significant (Study I: $r = .24$, $p < .05$; Study II: $r = .30$, $p < .01$). It was concluded that "attention deployment plays a significant role in RT performance, but it does not seem to be largely responsible for the relationship between intelligence and speed of information processing in the Hick paradigm" (Neubauer et al., 1992, p. 1331).

The Influence of Motivation

Marr and Sternberg (1987) consider individual differences in the motivation to perform quickly in tests of mental speed as an influential variable. Again, it is assumed that high-IQ subjects are more highly motivated in RT experiments than subjects of lower intelligence, thereby inflating the RT-IQ relationship. According to Detterman (1987), individual differences in motivation could be homogenized by introducing RT feedback

in the ECT; however, he did not specify a hypothesis about the effects on RT-intelligence correlations. According to Marr and Sternberg's argumentation, these correlations should decrease if individual differences in motivation are reduced. Following Detterman's suggestion, Neubauer et al. (1992) compared correlations between the Hick paradigm and intelligence from two groups (with vs. without RT feedback): RT feedback had the expected effects on Hick parameters (shorter RTs and smaller intraindividual variabilities), but the correlations between the Hick paradigm and psychometric intelligence did not differ significantly between the two groups. Contrary to Marr and Sternberg's conjecture, the RT feedback even produced a somewhat higher relationship between the bits/sec measure from the Hick paradigm and intelligence ($r = .37$) than was observable in the group without RT feedback ($r = .28$). Similar results have been reported by Lindley and Smith (1992) and by Larson, Saccuzzo and Brown (1994). Therefore the reduction of motivational variance seems to have the effect, if any, of reducing error variance, thereby allowing a better assessment of the true relationship between RTs and intelligence.

Timed vs. Untimed Intelligence Tests

Sternberg (1985) and others have suggested that the RT-intelligence relationship might be at least partly attributable to the fact that many psychometric intelligence tests are timed. Contrary to this presumption, Vernon, Nador and Kantor (1985) and Vernon and Kantor (1986) demonstrated that timed versus untimed administrations of intelligence tests produced rather similar correlations between RTs and intelligence. On the basis of Jensen's model (see the fourth section), Vernon et al. (1985) even expected higher RT-IQ correlations under untimed conditions: Only then did the subjects also attempt to answer the most complex test items, which placed the greatest demands on their information processing capabilities, that is, required more processing, storage and retrieval of information for item solution.

The question concerning the influence of timed vs. untimed intelligence on RT-IQ correlations can also be formulated the other way: Is there a relationship between SIP in ECTs and response latencies in psychometric IQ tests? The author (Neubauer, 1990b) attempted to answer this question by correlating RTs from the modified Hick paradigm (Neubauer, 1991) with response latencies assessed for each item of Raven's APM. As expected, Hick parameters correlated negatively with intelligence, but not with response latencies in the Raven matrices. This finding, too, contradicts the assumed inflation of RT-intelligence correlations through the use of timed intelligence tests.

MENTAL SPEED: ONE UNITARY OR SEVERAL INDEPENDENT PROCESSES?

In the previous section it was demonstrated that empirical tests of the major top-down explanations of the RT-intelligence relationship did not produce sound evidence for the presumed artifactual nature of this relationship (for more evidence see Neubauer, 1995). In light of these negative findings for top-down explanations and the considerable empirical evidence in favor of negative mental speed-intelligence correlations (see the second section), I conclude that a—presumably physiologically based—SIP constitutes an elemental basis of general intelligence.

Before pointing out one important unresolved issue let us first summarize the main empirical findings of the mental speed approach:

1. Empirical studies have provided sound evidence for a moderate, but significant negative relationship between RTs (and other parameters) in ECTs and psychometric intelligence. For the simpler ECTs, like the Hick and the IT paradigm, correlations between -.2 and -.3 have been observed. Somewhat higher correlations have been found for more complex ECTs like the Sternberg and Posner paradigms (up to -.4). The so-called complexity hypothesis has been advanced to explain these different figures.

2. Attempts to obtain a better IQ prediction by means of ECT batteries were motivated by Jensen's theoretical model, postulating not only SIP but also speed of STM and LTM retrieval as elementary determinants of complex problem-solving ability. As shown above, these attempts have met with mixed success.

3. The complexity hypothesis as well as the use of ECT-batteries are in at least partial contradiction to Jensen's assumption that only one property of the physiological substrate (the oscillation rate) should determine RT measures: Why should RT-intelligence correlations increase with ECT complexity and why should ECT-batteries provide a better IQ prediction if only one SIP determines the RT-IQ relationship? It must be this apparent contradiction that caused Jensen to renounce his assumption of one general SIP in a recent publication (Kranzler & Jensen, 1991). These authors questioned the "nature of psychometric *g*: Unitary process or a number independent process" and answered this question in favor of the second alternative by suggesting four independent SIP factors (derived by means of factor analysis from an ECT battery) as sources of *g*-variance.

In the following preliminary attempt shall be made to resolve this apparent contradiction (for a more detailed outline of these considerations see Neubauer, 1995). I will argue that at present a thorough study of empirical evidence does not warrant a renunciation of the "unitary *g*"-hypothesis.

For my argumentation, first, two kinds of ECT variables (from Hick, Sternberg, Posner paradigms and other ECTs) have to be distinguished, which I shall here call primary and secondary variables. Primary variables

are calculated directly from the RT trials (e.g., mean/median RT and SDRT) and are mostly indicative of processes common to all ECTs, like perception and discrimination of the stimulus, response-choice and the motoric response (following this argumentation the intercepts from the Hick and Sternberg paradigms also belong to the primary variables). Secondary variables (e.g., Sternberg slope and the NI-PI difference in the Posner paradigm) are derived from primary measures and are indicative of a process unique to the ECT, like STM access in the Sternberg paradigm and LTM retrieval in the Posner task.

Even primary variables from different ECTs are usually highly correlated, as they reflect common processes. To mention only a few studies: Keating and Bobbitt (1978) reported intercorrelations from .65 to .81; in a study by Jensen (1987b) *r*'s ranged from .38 to .57; a factor analysis of the high intercorrelations (mean *r*'s = .50 to .70) obtained in five samples studied by Vernon (cf. 1989 for a review) produced only one factor (with an eigenvalue above 1.0), which accounted for 65.5 to 83.4 percent of variance.

On the other hand, secondary variables usually display low intercorrelations: .19 to .35 in Keating and Bobbitt's study (1978); -.15 to -.30 in Jensen (1987b); in Vernon's 1983 study the intercorrelations of secondary variables (mean *r* = .16) gave three factors with eigenvalues above 1.0. Finally, correlations between primary and secondary ECT variables are low and both types of variables load on different factors when submitted to a factor analysis (McGue, Bouchard, Lykken, & Feuer, 1984).

We can, therefore, postulate two independent classes of processes in ECTs: First, the *basic processes* of stimulus perception and discrimination, response choice and motoric response, which are involved in most ECTs. Second, the *operational processes* like speed of access to STM and LTM. Now, let us see whether both types of processes provide (independent) sources of prediction of psychometric intelligence. Evidence already summarized clearly contradicts this assumption. Due to the multicollinearity (i.e., high intercorrelations) of primary ECT variables, Vernon (1983) as well as Jensen (1987b) found that the multiple *R*'s barely exceeded the largest zero-order correlations with psychometric intelligence. In both studies, these largest zero-order correlations were obtained with primary ECT variables. Secondary variables like the slope in the Sternberg paradigm and the NI-PI measure in the Posner paradigm did not produce a significant increment in multiple IQ predictions. Regarding the reviews of RT-intelligence correlations for the Sternberg and the Posner paradigms, this does not seem not surprising. Contrary to Jensen's expectation that the speed of STM access should also correlate with psychometric intelligence, the summary of studies correlating the Sternberg paradigm with psychometric intelligence (see Table 4) gave only a low mean correlation between slope and IQ (-.09 or -.11, *N*-weighted), whereas the primary or basic variables like mean/median RT, SDRT, but also the intercept gave *r*'s around

-.3. Therefore, the speed of STM retrieval does not account for substantial sources of IQ variance.

At first glance the findings for the Posner paradigm look more encouraging. The secondary measure from this ECT (the NI-PI measure, which should reflect LTM retrieval time) is substantially related to psychometric intelligence (mean r = -.29 or -.27, N-weighted, see Table 5). This mean correlation, however, should be viewed with caution, as three (out of five) of the NI-PI * IQ correlations in Table 5 are not with measures of general intelligence: Two studies used verbal measures, one study reports a correlation with "crystallized intelligence." From the remaining two studies, which used measures of general intelligence, one (McGue et al.) gave only a weak correlation of -.12.

Summarizing, the correlations between secondary variables from Sternberg and Posner paradigms as well as the rather unsuccessful attempts to obtain a better IQ prediction from multiple ECTs *do not support* Jensen's assumption that psychometric g is determined by a number of independent processes and is, therefore, not unitary. On the contrary, strong support for the "unity of g"-assumption can be inferred from a twin study by McGue and Bouchard (1989). Two outcomes are of special interest here:

1. Primary variables from the Sternberg, the Posner and a mental rotation paradigm displayed no patterns of discriminant validity with respect to four different cognitive ability domains; each of these variables correlated significantly negatively with at least three out of four cognitive abilities. On the other hand, the secondary variables (slopes from Sternberg and mental rotation paradigms, NI-PI in the Posner task) gave a clear picture of convergent and discriminant validities: The speed of STM and LTM access correlated only with verbal ability, the speed of mental rotations was associated only with spatial intelligence and a perceptual speed factor.

2. Primary ECT variables displayed substantial heritabilities (h^2 around .50). Yet the secondary variables turned out to be scarcely inherited (h^2-values around .20).

From these findings the authors conclude that primary ECT variables "may reflect aspects of the neurological structure ... that are inherited and that are important for performance on a wide range of cognitive tasks" (McGue & Bouchard, 1989, p. 41). In contrast, the secondary ECT measures "may primarily reflect alternative information-processing strategies that are acquired and not inherited and that are related to performance in specific cognitive ability domains" (McGue & Bouchard, 1989, p. 41). On the basis of the present state of knowledge, I would, therefore, conclude that a unitary process seems to be responsible for the relationship between psychometric intelligence and SIP.

IMPLICATIONS FOR APPLIED PSYCHOLOGY

Finally some words should be said about the practical utility of the mental speed approach. Some authors promised that this approach might provide us with new "intelligence tests" that—due to their apparent simplicity and minimal cognitive demands—should allow an almost culture-fair assessment of (biological) intelligence. However, the evidence presently available with RT-intelligence correlations mainly between -.2 and -.4 does not seem to warrant such a viewpoint. However, before refuting any practical usefulness of the mental speed approach on the basis of these at best moderate coefficients, one should bear in mind that correlations between different cognitive ability tests as well as validity coefficients of traditional intelligence tests rarely exceed such moderate degrees of association. The question should rather be put this way: Do ECTs allow a better prediction of external criteria than traditional IQ tests? Until now, only one attempt to assess the external validity of ECTs is known (to me). Kirby and Nettelbeck (1989) demonstrated that a combination of measures from the Hick paradigm allows a better discrimination (88.8% correct classifications) of three (extreme) groups (mental retardates vs. college vs. university students) than Raven's matrices (82.7%). Interestingly, an almost perfect discrimination (96.7%) resulted from the combination of both measures.

The claim of the mental speed approach to allow for a culture-fair assessment of some kind of biological intelligence also bears implications for educational psychology. A recent empirical attempt to relate learning ability to mental speed shall illustrate this. We know that learning ability is one of the most important correlates of human intelligence (Jensen, 1989). As the assessment of learning ability (and therefore of intelligence) by means of so-called learning tests (cf. Guthke, 1993) is also assumed to be less culturally biased and closer to the biological substrate, ECTs should correlate higher with such learning tests than with traditional IQ tests. Perhaps the recently reported empirical confirmation of this hypothesis (Guthke, Beckmann, & Stein, 1993) points to another practical application of the mental speed approach: The assessment of a person's learning ability or potential by means of elementary cognitive tasks.

NOTES

1. From Neubauer (1995).
2. From Neubauer (1995).
3. Readers familiar with the mental speed approach might miss A.E. and D.E. Hendricksons (1982) biochemical model here. In the author's view, however, this model has only historical value—if any—as it scarcely stimulated any research efforts, probably because of its "esoteric" (Deary & Caryl, 1993) nature.

REFERENCES

Ackerman, P. L., & Schneider, W. (1985). Individual differences in automatic and controlled information processing. In R. Dillon (Ed.), *Individual differences in cognition* (Vol. 2, pp. 35-66). New York: Academic.

Baumeister, A. A., & Kellas, G. (1968). Distribution of reaction times of retardates and normals. *American Journal of Mental Deficiency, 72*, 715-718.

Brickenkamp, R. (1978). *Test d2. Aufmerksamkeits-Belastungstest*. Göttingen: Hogrefe.

Carlson, J. S., & Widaman, K. F. (1987). Elementary cognitive correlates of *G*: Progress and prospects. In P. A. Vernon (Ed.), *Speed of information and intelligence* (pp. 69-99). Norwood, NJ: Ablex.

Ceci, S. J. (1990). On the relation between microlevel processing efficiency and macrolevel measures of intelligence—Some arguments against current reductionism. *Intelligence, 14*, 141-150.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3*, 472-517.

Cohn, S. J., Carlson, J. S., & Jensen, A. R. (1985). Speed of information processing in academically gifted youths. *Personality and Individual Differences, 6*, 621-629.

Deary, I. J. (1994). Sensory discrimination and intelligence—postmortem or resurrection. *American Journal of Psychology, 107*, 95-115.

Deary, I., & Caryl, P. (1993). Intelligence, EEG, and evoked potentials. In P. A. Vernon (Ed.), *Biological approaches to the study of human intelligence* (pp. 259-315). Norwood, NJ: Ablex.

Deary, I. J., Langan, S. J., Graham, K. S., Hepburn, D. & Frier, B. M. (1992). Recurrent severe hypoglycemia, intelligence, and speed of information processing. *Intelligence, 16*, 337-359.

Detterman, D. K. (1986). Human intelligence is a complex system of separate processes. In R. J. Sternberg & D. K. Detterman (Eds.), *What is intelligence*? (pp. 57-61). Norwood, NJ: Ablex.

Detterman, D. K. (1987). What does reaction time tell us about intelligence. In P. A. Vernon (Ed.), *Speed of information processing and intelligence* (pp. 177-200). Norwood, NJ: Ablex.

Dugas, J. L., & Kellas, G. (1974). Encoding and retrieval processes in normal children and retarded adolescents. *Journal of Experimental Child Psychology, 17*, 177-185.

Galton, F. (1883). *Inquiries into human faculty and it's development*. London: Macmillan.

Goldberg, R. A., Schwartz, S., & Stewart, M. (1977). Individual differences in cognitive processes. *Journal of Educational Psychology, 69*, 9-14.

Guthke, J. (1993). Developments in learning potential assessment. In J. H. M. Hamers, K. Sijtsma & A. J. J. M. Ruijssenaars (Eds.), *Learning potential assessment: Theoretical, methodological and practical issues*. Amsterdam: Swets & Zeitlinger.

Guthke, J., Beckmann, J. F., & Stein, H. (1993). *Recent research evidence on the validity of learning tests*. Paper presented at the 3rd European Congress of Psychology. Tampere, Finland.

Harris, G. J., & Fleer, R. E. (1974). High speed memory scanning in mental retardates: Evidence for a central processing deficit. *Journal of Experimental Child Psychology, 17*, 452-459.

Hemmelgarn, T. E., & Kehle, T. J. (1984). The relationship between reaction time and intelligence in children. *School Psychology International, 5*, 77-84.

Hendrickson, A. E. (1982). The biological basis of intelligence Part I: Theory. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 151-196). Berlin: Springer.

Hendrickson, D. E. (1982). The biological basis of intelligence Part II: Measurement. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 197-228). Berlin: Springer.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology, 4*, 11-26.

Hunt, E. (1980). Intelligence as an information processing concept. *British Journal of Psychology, 71*, 449-474.

Hunt, E., Davidson, J., & Lansman, M. (1981). Individual differences in long-term memory access. *Memory and Cognition, 9*, 599-608.

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

Jenkinson, J. C. (1983). Is speed of information processing related to fluid or to crystallized intelligence? *Intelligence, 7*, 91-106.

Jensen, A. R. (1982). Reaction time and psychometric *g*. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 93-132). Berlin: Springer.

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

Jensen, A. R. (1987b). Process differences and individual differences in some cognitive tasks. *Intelligence, 11*, 107-136.

Jensen, A. R. (1989). The relationship between learning and intelligence. *Learning and Individual Differences, 1*, 37-62.

Jensen, A. R. (1992). Understanding *g* in terms of information processing. *Educational Psychology Review, 4*, 271-308.

Jensen, A. R., & Munro, E. (1979). Reaction time, movement time, and intelligence. *Intelligence, 3*, 121-126.

Keating, D. P., & Bobbitt, B. L. (1978). Individual and developmental differences in cognitive-processing components of mental ability. *Child Development, 49*, 155-167.

Kirby, N. H. & Nettelbeck, T. (1989). Reaction time and inspection time as measures of intellectual ability. *Personality and Individual Differences, 10*, 11-14.

Kranzler, J. H. (1991). Reaction time, response errors, and *g*. *Perceptual and Motor Skills, 73*, 1035-1043.

Kranzler, J. H., & Jensen, A. R. (1989). Inspection time and intelligence: A meta-analysis. *Intelligence, 13*, 329-348.

Kranzler, J. H., & Jensen, A. R. (1991). The nature of psychometric *g*—unitary process or a number of independent processes. *Intelligence, 15*, 397-422.

Kranzler, J. H., Whang, P. A., & Jensen, A. R. (1988). Jensen's use of the Hick paradigm: visual attention and order effects. *Intelligence, 12*, 379-392.

Lansman, M., Donaldson, G., Hunt, E. & Yantis, S. (1982). Ability factors and cognitive processes. *Intelligence, 6*, 347-386.

Larson, G. E., Merritt, C. R., & Williams, S. E. (1988). Information processing and intelligence: Some implications of task complexity. *Intelligence, 12*, 131-148.

Larson, G. E., & Saccuzzo, D. P. (1986). Jensen's reaction-time experiments: Another look. *Intelligence, 10*, 231-238.

Larson, G. E., Saccuzzo, D. P., & Brown, J. (1994). Motivation—cause or confound in information processing intelligence correlations. *Acta Psychologica, 85*, 25-37.

Lemmon, V. W. (1927). The relation of reaction time to measures of intelligence, memory and learning. *Archives of Psychology, 15*, 5-38.

Levine, G., Preddy, D., & Thorndike, R. L. (1987). Speed of information processing and level of cognitive ability. *Personality and Individual Differences, 8*, 599-608.

Lindley, R. H., Bathurst, K., Smith, W. R., & Wilson, S. M. (1993). Hick's law, IQ, and singularity or specificity of mind—a psychometric analysis. *Personality and Individual Differences, 15*, 129-135.

Lindley, R. H., & Smith, W. R. (1992). Coding tests as measures of IQ: Cognition or motivation. *Personality and Individual Differences, 13*, 25-29.

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

Longstreth, L. E. (1986). The real and the unreal: A reply to Jensen and Vernon. *Intelligence, 10*, 181-191.

Marr, D. B., & Sternberg, R. J. (1987). The role of mental speed in intelligence: A triarchic perspective. In P. A. Vernon (Ed.), *Speed of information processing and intelligence* (pp. 271-294). Norwood, NJ: Ablex.

Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the radex and hierarchic models of intelligence. *Intelligence, 7*, 107-127.

McCauley, C., Dugas, J., Kellas, G., & DeVellis, R. F. (1976). Effects of serial rehearsal training on memory search. *Journal of Educational Psychology, 68*, 474-481.

McGue, M., Bouchard, T. J., Lykken, D. T., & Feuer, D. (1984). Information processing abilities in twins reared apart. *Intelligence, 8*, 239-258.

McGue, M., & Bouchard, T. J. Jr. (1989). Genetic and environmental determinants of information processing and special mental abilities: A twin analysis. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 5, pp. 7-45). Hillsdale, NJ: Erlbaum.

Miller, L. T., & Vernon, P. A. (1992). The general factor in short-term memory, intelligence, and reaction time. *Intelligence, 16*, 5-29.

Neubauer, A. C. (1990a). Selective reaction times and intelligence. *Intelligence, 14*, 79-96.

Neubauer, A. C. (1990b). Speed of information processing in the Hick paradigm and response latencies in a psychometric intelligence test. *Personality and Individual Differences, 11*, 147-152.

Neubauer, A. C. (1991). Intelligence and RT: A modified Hick paradigm and a new RT paradigm. *Intelligence, 15*, 175-193.

Neubauer, A. C. (1994). *Intelligenz und Geschwindigkeit der Informationsverarbeitung* [Intelligence and speed of information processing]. Graz: Unpublished habilitation thesis at the Karl-Franzens-University of Graz.

Neubauer, A. C.(1995). *Intelligenz und Geschwindigkeit der Informationsverarbeitung* [Intelligence and speed of information processing]. Vienna, New York: Springer.

Neubauer, A. C., Bauer, C., & Hoeller, G. (1992). Intelligence, attention, motivation and speed-accuracy trade-off in the Hick paradigm. *Personality and Individual Differences, 13*, 1325-1332.

Neubauer, A. C., & Freudenthaler, H. H. (1994). Reaction times in a sentence-picture verification test and intelligence: Individual strategies and effects of extended practice. *Intelligence, 19*, 193-218.

Neubauer, A. C., & Bucik, V. (1996). The mental speed-IQ relationship: Unitary or modular. *Intelligence, 22*, 23-48.

Posner, M. I., & Mitchell, R. F. (1967). Chronometric analysis of classification. *Psychological Review, 74*, 392-409.

Raven, J. C. (1958). *Advanced progressive matrices*. London: Lewis.

Raven, J. C. (1960). *Standard progressive matrices*. London: Lewis.

Roth, E. (1964). Die Geschwindigkeit der Verarbeitung von Information und ihr Zusammenhang mit Intelligenz [The speed of information processing and its relationship to intelligence]. *Zeitschrift für experimentelle und angewandte Psychologie, 11*, 616-622.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*, 1-66.

Sharp, S. (1898-1899). Individual psychology: A study in psychological method. *American Journal of Psychology, 10*, 329-391.

Smith, G. A., & Stanley, G. (1987). Comparing subtest profiles of *g* loadings and correlations with RT measures. *Intelligence, 11*, 291-298.

Spearman, C. (1927). *The Abilities of Man*. New York: Macmillan.

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

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist, 57*, 421-457.

Vernon, P. A. (1983). Speed of information processing and general intelligence. *Intelligence, 7*, 53-70.

Vernon, P. A. (1989). The generality of *g*. *Personality and Individual Differences, 10*, 803-804.

Vernon, P. A. (1993). Intelligence and neural efficiency. In D. K. Detterman (Ed.), *Current topics in human intelligence* (Vol. 3, pp. 171-187). Norwood, NJ: Ablex.

Vernon, P. A., & Jensen, A. R. (1984). Individual and group differences in intelligence and speed of information processing. *Personality and Individual Differences, 5*, 411-423.

Vernon, P. A., & Kantor, L. (1986). Reaction time correlations with intelligence test scores obtained under either timed or untimed conditions. *Intelligence, 10*, 315-330.

Vernon, P. A., Nador, S., & Kantor, L. (1985). Reaction times and speed of information processing: Their relationship to timed and untimed measures of intelligence. *Intelligence, 9*, 357-374.

Vernon, P. A., & Weese, S. E. (1993). Predicting intelligence with multiple speed of information-processing tests. *Personality and Individual Differences, 14*, 413-419.

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

Welford, A. T. (1986). Longstreth versus Jensen and Vernon on reaction time and IQ: An outsider's view. *Intelligence, 10*, 193-195.

Westhoff, K., & Kluck, M.-L. (1983). Zusammenhang zwischen Intelligenz und Konzentration [Correlation between intelligence and concentration]. *Diagnostica, 29*, 310-319.

Widaman, K. F., & Carlson, J. S. (1989). Procedural effects on performance on the Hick paradigm: Bias in reaction time and movement time parameters. *Intelligence, 13*, 63-85.

Wissler, C. (1901). The correlation of mental and physical tests. *Psychological Review Monographs, 3*, whole No 16.

INTELLIGENCE, PERSONALITY AND INFORMATION-PROCESSING:

AN ADAPTIVE PERSPECTIVE

Gerald Matthews

INTRODUCTION

There is an extensive literature showing that intelligence test scores predict performance on a wide range of information processing tasks. However, the wealth of data now available is in danger of hindering efforts to integrate the cognitive psychology of intelligence into wider intelligence theory. Currently, there is no theory of intelligence and information-processing which can explain the full patterning of cognitive correlates of intelligence, including the near-independence of some processing functions from intelligence (Matthews & Dorn, 1995). This problem is not confined to intelligence research. Cognitive psychological research on stress and on personality has similarly resulted in a complex and somewhat arbitrary list of processing functions related to stress manipulations and to personality traits, which resist parsimonious theoretical interpretation (e.g., Hockey, 1984; Matthews, 1992). Personality data are

Advances in Cognition and Educational Practice, Volume 4, pages 175-200.

ISBN: 0-7623-0105-8

particularly confusing, in that, although intelligence is psychometrically distinct from major personality dimensions such as extraversion and neuroticism (e.g., Saklofske & Kostura, 1990), intelligence and personality influence some common processing functions. For example, short term memory (STM) varies with intelligence, extraversion and neuroticism (Matthews & Dorn, 1995). This conflict between distinctiveness at a psychometric level of description, and overlap at the information processing level is difficult for cognitive theory to accommodate.

In this chapter, I will argue that, although cognitive psychological approaches are indispensable for a meaningful characterization of individual differences, cognitive theory does not on its own provide a satisfactory explanation for the patternings of processing functions associated with individual difference factors. Instead, the cognitive level of description must be supplemented by adaptational explanations. It is not sufficient to demonstrate that, say, intelligence correlates with speed of reaction, and then to assume that fast reactions are a central component of intelligent behavior. It is necessary to determine what categories of real-world behavior are enhanced by higher intelligence, and investigate the contribution that speed of reaction makes to those behaviors (if any). It will be claimed that intelligence and personality factors differ in their adaptational significance, and their common information-processing correlates support differing adaptive skills.

The structure of the chapter is follows. First, the adaptational function of intelligence is considered—we must ask "what is intelligence for?" I will claim that, in contemporary industrial societies, intelligent individuals have a behavioral advantage in a restricted set of environments characterized by complexity and novelty. Next, the information processing correlates of intelligence will be briefly reviewed, and their relationship to adaptation assessed. I will then review how two major personality traits, extraversion and neuroticism, relate to information-processing, and distinguish their adaptational significance from that of intelligence. The chapter has some limitations in scope. I will assume that general intelligence (*g*) measured by psychometric tests is a psychologically meaningful construct. I will not be much concerned with primary abilities, which may represent somewhat more specialised intellectual adaptations than *g*, because most (but not all) of the predictive validity of intelligence measures derives from general intelligence (e.g., Hunter & Hunter, 1984). I will also neglect issues related to cross-cultural research, such as use of psychometric tests in non-technological societies.

ADAPTIVE FUNCTIONS OF GENERAL INTELLIGENCE

Biological and Psychological Conceptions of Adaptation

It is important from the outset to distinguish between- and within-species adaptive functions. There has been considerable informed speculation on the selection pressures which lead to the evolution of such unique intellectual capabilities of the human species as language, reasoning and advanced problem-solving. Behaviors such as tool use, cooperative predation and food-gathering, and shelter construction have all been linked to the evolution of intelligence (Richards, 1987). There are considerable problems in establishing causality. On the one hand, the acquisition by the human species of a new species-specific behavior (for whatever reason) may increase the adaptive value of high intelligence. Conversely, intelligence increases the likelihood of selection for new, intellectually sophisticated behaviors, or the learning of new behaviors which may then be culturally transmitted through the generations which follow.

However, as Jerison (1982) points out, within-species variability in intelligence may not be closely linked to between-species variability. That is, the psychological characteristics which distinguish more and less intelligent humans may not correspond to those characteristics which distinguish the human species from other animal species. Hence, the primary source of data for establishing the contemporary utility of intelligence is studies of important real-life behaviors and other criteria. There is a further problem here in establishing which behaviors are important, and whether their outcome is successful to the individual. From the evolutionary biological perspective the issue is the role of behavior in increasing likelihood of genetic fitness and selection, but relationships of this kind are difficult to establish in studies of humans. In research practice, cultural definitions of success, such as occupational success, are usually taken as benchmarks for adaptation. I will follow this convention in asking how intelligence is beneficial within advanced societies, leaving open the issue of how socially-valued attainments relate to Darwinian fitness. My primary concern is with phenotypic rather than genotypic intelligence, and with behavioral rather than evolutionary success.

Contextual Perspectives on Intelligence

In this restricted sense of adaptation, the most thorough account of the adaptive relevance of intelligence is that provided by Sternberg's (1985) triarchic theory of intelligence. Sternberg proposes a contextual perspective, within which intelligence is defined as "purposive adaptation to, and selection and shaping of, real-world environments relevant to one's life" (Sternberg, 1985, p. 45). Intelligence refers not just to efficient solving of real-world "problems" but to active attempts to select and manipulate environments.

However, as Sternberg (1985) acknowledges, there is rather little direct evidence on this role for intelligence, in part because "practical intelligence" is difficult to assess. Studies of people's implicit theories of intelligence (Sternberg, Conway, Ketron, & Bernstein, 1981) indicate the nature of socially-valued skills; Sternberg et al. (1981) identified three main constellations of skills, described as practical problem-solving ability, verbal ability and social competence. It is plausible that *g* contributes to skills within the first two constellations, but social intelligence has proved a problematic concept which has not been satisfactorily operationalized (Brody, 1992). Sternberg and Wagner (1989) discuss measures of the "tacit knowledge" required to succeed in various real-world contexts, which are largely independent of *g*, but predict certain criteria of success in business and academic occupations more strongly than do measures of *g*.

A more radical contextualist theory of intelligence and adaptation is that of Ceci (1990). He rejects the existence of a substantial general ability factor, and argues instead for a set of multiple knowledge domains whose adaptive significance is highly context-dependent. In the present context, in which no challenge is made to the importance of general intelligence, the most interesting feature of Ceci's (1990) book is the evidence it presents on *g* and certain real-world skills. He describes several empirical studies in which psychometric measures of general intelligence fail to predict level of skill on complex tasks, such as filling complex orders to a dairy, and handicapping horse races. A more dramatic example of the limitations of *g* as a predictor of real-world performance is instances of "idiots savants." Howe (1989) documents various cases in which people of very low measured intelligence show mastery of complex skills such as mental arithmetic and calendar-date calculations.

Such results do not in themselves invalidate the concept of general intelligence, which is quite compatible with low *g* individuals developing certain complex skills, perhaps through obsessional practice (Nettelbeck, 1990). The cautious inference to be drawn from the various findings just discussed is that intelligence is more predictive of some real-world criteria than of others. In fact, the real-world criterion most strongly linked to intelligence is educational achievement, which typically correlates at about .4 to .6 with IQ or *g* measures (Brody, 1992; Snow & Yalow, 1982). Effects of *g* on educational achievement are not an artifact of social class effects, and social class has little effect on achievement when *g* and other mediating variables are controlled (Rehberg & Rosenthal, 1978). *g* is also quite strongly related to job performance; Hunter and Hunter's (1984) review of the area indicates a mean corrected correlation between *g* and supervisor-rated performance of .53. The importance of *g* increases with job complexity. However, good job performance does not necessarily receive its due rewards, and studies investigating occupational status tend to de-emphasise the role of *g* to some extent. Effects of *g* on occupational status appear to be primarily indirect; *g* influences education which is the main

influence on status (Duncan, Featherman & Duncan, 1972). It is also uncertain whether *g* has any direct effect on income, with different studies arriving at different results (Ceci, 1990; Jencks, 1979).

Intelligence, Novelty and Transfer of Skills

It appears that intelligence may only be strongly adaptive in a restricted set of real-world contexts requiring the intellectual skills associated with higher *g*. Studies of global real-world success criteria are not, however, particularly informative concerning the nature of the skills which support correlations between *g* and achievement. Theories of intelligence have addressed this issue. Sternberg (1985) suggests that intelligence is particularly characterized by ability to deal with novel tasks and situations, and to automatise performance. The former proposition is supported by evidence that intelligence is related to learning particularly when the task requires insight, the acquisition of a systematic set of skills or knowledge, and when the learner must decide how to transfer existing knowledge to a new skill (Jensen, 1980). Similarly, work on mental retardation (e.g., Ferretti & Cavalier, 1991) identifies transfer of strategies to new problems, in the absence of explicit instructions, as a key feature of intelligence.

Clearly, conceptual learning and transfer abilities are likely to contribute particularly to educational success. Intelligence may have relatively weak direct effects on occupational success because learning of new intellectual skills is less important in most jobs than it is in education. In addition, of course, the more intellectually demanding the job, the higher the degree of initial selection for intelligence, so that correlations between intelligence and attainment are likely to be highly attenuated. Studies of aptitude-treatment interaction (ATI) (Cronbach & Snow, 1977; Snow & Yalow, 1982) demonstrate that intelligence is more adaptive for certain learning environments than others. Although the positive IQ—achievement correlation is ubiquitous, the size of the correlation varies with instructional treatment. More intelligent learners are particularly advantaged when information-processing load is high and unstructured, and learners must build their own cognitive structures as the intellectual skill is acquired (Cronbach & Snow, 1977). In contrast, intelligence is not predictive of rote learning ability, though few contemporary instructional regimes rely exclusively on rote learning.

Thus, I shall take as a working hypothesis the view that individual differences in intelligence are most adaptive in real-world environments characterised by novelty, unpredictability and variability, such that the individual must develop new skills dependent on systematic knowledge over time. I submit that neither Ceci's (1990) low IQ expert handicappers or Howe's (1989) idiots savants would function well in such environments, but the flexibility of intellectual skill of high IQ individuals would provide an adaptive advantage. Of course, the range

of important contexts in real-life which can be characterised as having such properties is moot; many people are inclined to avoid complex and unpredictable situations.

INTELLIGENCE AND INFORMATION-PROCESSING

In this section, I provide a brief review of the cognitive correlates of general intelligence, and consider their adaptive function. More detailed reviews are provided by Brody (1992), Larson and Alderton (1992) and Matthews and Dorn (1995). It is convenient in this context to divide information-processing tasks into four categories: *elementary information-processing* (EIP) tasks, *attentional* tasks, *working memory* tasks and *skilled performance* tasks.

Speed of Elementary Information-Processing

EIP tasks require a minimum of internal computation; performance depends on speed and accuracy of simple perceptual discriminations, and on speed of initiation of simple responses. Simple and choice reaction time (RT) tasks are paradigmatic examples. Correlations between *g* and such tasks are well-documented. Jensen (1987) reviews studies of RT which show that the mean correlation with *g* for simple RT is -.18, whereas for choice RT it rises to -.23. Another EIP task correlated with *g* is inspection time (IT). On this simple perceptual task, the subject must decide which of two masked lines is the longer (Brand & Deary, 1982). Kranzler and Jensen (1989) calculated a mean correlation of -.29 between *g* and IT, on the basis of data from a number of published studies. Empirical data may under-estimate the strength of the underlying association between *g* and IT due to statistical factors such as unreliability of variables and restriction of range of scores. The corrected correlation between performance IQ and IT may be as much as -.55 (Kranzler & Jensen, 1989).

The evidence available points towards a genuine association between *g* and speed of certain elementary processes. Attempts have been made to argue that both *g* and EIP measures are confounded by more complex processes such as strategy use and maintenance of focused attention (e.g., Longstreth, 1984). However, studies which have controlled for strategy use (Egan, 1994; Neubauer, 1991), or which have manipulated information-processing demands of RT tasks (Matthews & Dorn, 1989), have failed to confirm this general hypothesis. One unresolved issue is whether EIP can be characterised by a single speed-of-processing parameter, or whether there are distinct categories of elementary processing which may relate to different speed factors. Some researchers have distinguished psychometrically distinct processing functions; Nettelbeck and Rabbit (1992) showed that RT and IT measures make independent contributions to the prediction of *g*.

Attentional Resources and Control

Attentional tasks are those requiring selection between stimuli or attributes of stimuli, or direction of intensive processing effort towards particular stimuli. Most research effort has been directed towards testing relationships between *g* and measures of intensiveness of attention. *g* is weakly related or even unrelated to performance in some attentional paradigms, such as vigilance (Davies et al., 1984) and visual focused attention (Blanco & Alvarez, 1994). However, there is convincing evidence that *g* is often more strongly related to dual- than to single-task performance (e.g., Roberts, Beh, & Stankov, 1988). At a theoretical level, an attractive idea, dating to back to Spearman (1923), is that *g* relates to the energization of performance. In contemporary research, performance energetics are often formalised through the concept of attentional resources (Norman & Bobrow, 1975). Resources are a metaphorical pool of energy or fuel for processing which can be allocated flexibly to certain processing operations. Performance deteriorates when capacity of the energy reservoir is insufficient to meet the total demands for resources imposed by current processing. Adding an additional task is the most usual way to induce resource insufficiency experimentally, so the association between *g* and dual-task performance is suggestive of greater resource availability in more intelligent individuals.

However, a careful program of research conducted by Stankov (e.g., 1989), Spilsbury (1992) and others has provided evidence inconsistent with this hypothesis. Instead, it appears that the critical factor moderating *g*-performance correlations in dual-task performance is not resource requirements but task complexity. More intelligent individuals appear to be better at certain of the specific executive processing operations required to combine performance of two concurrent tasks (Matthews & Dorn, 1995; Spilsbury, 1992). However, intelligence probably does not relate to *all* executive functions related to divided attention. Intelligence may not be related to task-specific time-sharing abilities predictive of efficient dual-task performance (Brookings, 1990) or to co-ordination of multiple task components (Yee, Laden, & Hunt, 1994). Moreover, intelligence is not generally related to optimality of strategy use in attentionally-demanding tasks (Alderton & Larson, 1994). More research is required to identify the specific executive processes related to intelligence. Duncan (in press) presents data suggesting that a critical executive function may be keeping track of multiple goals during novel behavior; on a spatial attention task, less intelligent individuals tended to neglect a cue to move the spatial focus of attention when other goals were active, even though they understood the significance of the cue.

Working Memory

Working memory tasks require integration of short-term storage and internal computation; mental arithmetic is a typical task with high working memory requirements. The association between *g* and measures of short-term retention such as digit span is well-known. Research in this area has received fresh impetus from cognitive psychological research which suggests that working memory may be supported by several distinct processing systems (Baddeley, 1986). Matthews and Dorn (1995) derive two main conclusions from research on *g* and components of working memory. First, the strongest correlations between *g* and performance are obtained when the working memory task is complex, and may depend on a variety of sub-systems (e.g., Kyllonen & Christal, 1990). Measures of individual sub-systems are more weakly related to *g*, and some aspects of working memory such as use of sub-vocal articulation may not relate to *g* at all (Cantor, Engle, & Hamilton, 1991). Matthews and Dorn (1995) point out that *g*-working memory associations may, in part, be dependent on the same mechanisms contributing to correlations between *g* and speed of EIP, and between *g* and dual-task performance. Dempster (1981) has shown that digit span relates to individual differences in item identification, an elementary process. It is also likely that complex working memory tasks, such as those used by Kyllonen and Christal (1990), require executive control processes for co-ordinating multiple component processes, similar to those required for dual-task performance.

Skilled Performance

Skilled performance tasks require sequencing and integration of multiple processing routines; both stimuli and responses may be complex. Much of the empirical data in this area comes from applied studies of ability and performance of occupationally relevant tasks. While the predictive validity of ability tests in the occupational and educational arenas is no longer controversial (e.g., Hunter & Hunter, 1984; Snow & Yalow, 1982), the theoretical significance of applied studies is often unclear. Construct validity of both ability measures and performance criteria is often uncertain. The most fully articulated theory of intelligence and skill is that of Ackerman (1988, 1992). He proposes that ability—performance correlations vary with both level of practice and *consistency of the mapping* of the task. Consistency of mapping refers to the consistency with which given task stimuli evoke the same response on different occasions. General intelligence is most strongly related to performance early in learning, and when S-R mappings are inconsistent, as demonstrated empirically in studies of semantic category search and air-traffic control. Overlearnt cognitive-motor skills such as those required for simple

psychomotor performance are often largely independent of intelligence (Fleishman & Quaintance, 1984).

As with simpler tasks, the key function related to intelligence may be certain of the executive processes co-ordinating the operation of more basic processes. For example, Duncan, Williams, Nimmo-Smith and Brown (1992) showed that the driving skills most vulnerable to dual-task interference were also those most highly correlated with intelligence (although correlations were generally weak in this study). Snow (1989) reviews studies of intelligence and learning which suggest that high *g* learners are particularly advantaged in unstructured learning environments, because of their flexibility in developing personalised strategies without explicit instruction. Similarly, using a complex economic management simulation, Shute, Glaser and Raghavan (1989) showed that more intelligent individuals were more successful in planning "experiments" which increased understanding of the behavior of the simulation. Hence, although *g* does not always predict adaptive strategy use in experimental studies, it may be an important correlate of strategic processing when strategies are complex or involve multiple goals.

Intelligence and Information-Processing: Conclusions

Some cognitive processing operations are more strongly related to psychometric intelligence than others. Although the wealth of empirical evidence resists simple conclusions, two types of correlate of *g* stand out. The first is speed of EIP which is consistently related to *g* across a range of rather different tasks. The second is the efficiency of certain executive control processes, contributing to performance of multiple concurrent tasks, complex working memory tasks, and performance of skills characterised by variability of S-R mapping and limited practice. Conversely, many of the "stock-in-trade" constructs of the cognitive psychologist, such as attentional resources and capacity of specific stores in STM, seem to be, at most, weakly related to *g*. It is somewhat paradoxical that *g* should relate most strongly to both elementary and high-level processes, and more weakly to processes at an intermediate level of complexity and abstraction. A possible resolution to the paradox is that development of high-level executive processes is constrained by speed of elementary information processing (cf. Necka, 1992). Anderson (1992) suggests that a speed of a "basic processing mechanism" limits the complexity of symbolic knowledge acquisition routines, and, hence, general intelligence.

ADAPTIVE SIGNIFICANCE OF COGNITIVE CORRELATES OF INTELLIGENCE

Limitations of Information-Processing Measures

The stereotype of the brilliant but absent-minded professor who cannot function effectively in the real world is familiar (though possibly unfair!). Intelligent individuals may shine in the laboratory, but ability on information-processing tasks does not necessarily transfer to significant real-world skills. Furthermore, even if intelligence predicts skill, skill does not necessarily predict adaptive outcome. Consider the example of vehicle driving, for example, which is of interest as a set of skills which directly relate to survival and mortality. Motor vehicle accidents, which are frequently caused by human error, are a major cause of death in developed societies. The cognitive correlates of intelligence just described would suggest an adaptive advantage for the intelligent driver, due to faster reactions and to superior multiple-task performance. In fact, as meta-analysis of predictors of accident rates shows (Arthur, Barrett, & Alexander, 1991), neither speed of reaction nor intelligence confers much advantage. In the meta-analysis, the correlation between RT and accident involvement was only .053 ($N = 5{,}406$), whereas the correlation for psychometric ability tests was a modest .117 ($N = 1{,}020$). The best cognitive predictor of accident rate was a measure of flexibility of selective attention ($r = .257$, $N = 1{,}101$), related to efficiency of switching between channels of attention, a relatively specific attentional skill whose relationship to IQ is uncertain. Thus, although it is easy to envisage circumstances in which speed of reaction and other IQ-correlated processes might contribute to survival, the implication of the accident data is that other factors may be considerably more important in practice. More generally, there is a lack of evidence on whether fast reactions are adaptive in everyday life.

Skills as Links Between Information-Processing and Adaptation

Measures of individual processing functions are probably too far removed from most real-life adaptive behaviors to be directly relevant to the adaptation issue. What is required is a skills-based analysis, in which processing functions are seen as contributing to specific skills, which are the more direct influence on success of adaptation. For example, there has been considerable research on the contribution of components of working memory to various verbal skills such as reading and comprehension, which typically correlate with general or verbal ability (e.g., Daneman & Carpenter, 1980). We have seen that it may be the central executive component of working memory which is most strongly related to *g* (Kyllonen & Chrystal, 1990). In a review of studies of working memory and language, Gathercole and Baddeley (1993) suggest that the best-

established function of the central executive is its contribution to comprehension, through more efficient processing of semantic and syntactic information. Monitoring and coordination of use of multiple processing resources is another executive function which may relate to individual differences in both intelligence and comprehension (Swanson, 1993). The central executive may also aid vocabulary acquisition, again through semantic processing, and speech production, through planning the conceptual content of speech (Gathercole & Baddeley, 1993). Modelling of data sets in this area is usually unsophisticated, but it is plausible that the superior central executive functioning associated with higher intelligence leads to adaptive advantages in a number of important verbal skills. However, other cognitive processes which are substantially independent from intelligence may also contribute to verbal skills. Gathercole and Baddeley (1993) review evidence that individual differences in short-term retention based upon sub-vocal articulation predict development of skills such as comprehension, vocabulary acquisition and reading even when intelligence is controlled.

As a second example, various ability factors appear to predict performance of computer-based tasks such as word-processing (Czaja, Joyce, & Hammond, 1989; Egan & Gomez, 1985) and data-base search (Vicente & Williges, 1988). Westerman, Davies, Glendon, Stammers and Matthews (in preparation) showed that: (1) word-processing performance, (2) component skills contributing to word-processing and (3) ability measures were mutually inter-correlated. It appeared that skills which correlate with ability, such as locating text in the display, directly contributed to overall performance. As in the case of language skills, processing functions correlated with intelligence may directly contribute to computerised task skills, but these functions may not fully explain individual differences in overall skill. It is plausible that adaptive benefits accrue from specific language skills and from aptitude for computerised tasks such as word-processing, although direct evidence is rather lacking.

Characterizing Real-World Skills Associated With Intelligence

It appears that both intelligence and some of the processing functions with which it correlates predict some real-world skills, and possibly adaptive outcomes, but the data available provide only limited indications of when intelligence is most adaptive. The implication of Ackerman's (1988) analysis of individual differences is that intelligence is adaptive primarily when successful action requires complex, and possibly relatively unpractised, skills, as in the examples of comprehension of novel verbal stimuli and acquiring skills related to new technology. This characterization of intelligence is consistent with Sternberg's (1985) view of dealing with novelty as a key aspect of intelligence, and with the hypothesis that intelligence relates to flexibility

of application of existing skills by modifying them to fit new tasks and circumstances (Campione, Brown, & Ferrara, 1982).

It may be inferred that the benefits of intelligence for individuals depend on the environmental and social structures within which they are embedded. Contemporary industrial societies characterized by rapid technological and social change may particularly favor the more intelligent person. Skills for dealing with the physical world rapidly become obsolete and must be replaced with new skills, such as replacing typewriting skills with word-processing skills. In "primitive" non-technological societies, in which a fixed set of skills are transmitted from generation to generation, intelligence may be of reduced adaptive value. It is also possible that the verbal skills associated with intelligence are more adaptive when social norms vary over time, as they do in Western societies, as exemplified by changes in attitudes towards women and ethnic minorities during the last 30 years. In stable societies, the skills required for management of social interactions may be relatively predictable and easy to learn, or even codified explicitly, as in Victorian etiquette books. In contemporary society, encounters with others, perhaps of differing cultural background, may require a greater degree of improvization, increasing the advantage associated with higher intelligence.

EXTRAVERSION, PERFORMANCE AND ADAPTATION

There is an increasing consensus that are five or six major dimensions of personality. The "Big Five" model (Costa & McCrae, 1992) posits dimensions of Extraversion, Neuroticism, Conscientiousness, Agreeableness and Openness. There are some unresolved issues, such as the intriguing possibility that intelligence itself should be seen as a personality dimension, within a "Comprehensive Six" framework (Brand, 1994). Matthews and Oddy (1993) draw attention to uncertainties over the status of traits related to ambition and activity within the Big Five. Nevertheless, the basic psychometric framework for future personality research is becoming increasingly clear (Deary & Matthews, 1993). Research on information-processing and performance has largely concerned two broad traits; extraversion and neuroticism. (Extraversion is sometimes investigated in the guise of its impulsivity component: e.g., Humphreys & Revelle, 1984). Next, I present a brief review of the information-processing correlates of extraversion, and their adaptive significance. More detailed reviews of extraversion and performance are provided by Eysenck and Eysenck (1985), Matthews (1992) and Matthews and Dorn (1995). Furnham (1992) provides a comprehensive survey of occupational criteria related to personality.

Information-Processing Correlates of Extraversion

Extraversion has at least two distinct kinds of effect on performance. First, it has main effects; extravert—introvert differences which are apparent in the absence of any manipulation of the environment. Second, extraversion moderates the effects on performance of motivationally significant contextual variables, and of emotional states. Main effects of extraversion on EIP tasks are inconsistent and often non-significant (e.g., Amelang & Ullwer, 1991). Where effects are found, extraverts typically show faster speed of response, often accompanied by an increased error rate (Eysenck & Eysenck, 1985). Main effects of extraversion are more pronounced on demanding attentional and short-term memory tasks, on which extraverts normally show superior performance to introverts, particularly when stimuli are verbal in nature (Matthews, 1992). Eysenck (1982) characterizes much of the evidence as suggesting that extraverts have more attentional resources available, leading to better resistance to distraction, dual-task performance and working memory.

In contrast to this general trend, introverts are reliably superior on vigilance tasks (Koelega, 1992). Koelega's (1992) meta-analysis also shows that extraverts tend to adopt a lower response criterion; that is, they are readier to respond than introverts, a point also made by Brebner and Cooper (1985). The apparent paradox, that extraverts have more attentional resources, but poorer sustained attention, is resolved by considering the processing demands of sustained attention more closely. In fact, it seems that only somewhat atypical tasks, made highly demanding by factors such as high event rate, degraded stimuli, memory load and so forth, are strongly resource-limited (Matthews, Davies, & Holley, 1993). Performance of traditional, low-event rate tasks appears to be limited by other factors, such as expectancy and tolerance of low arousal (Davies & Parasuraman, 1982). There is something of a trend towards superior performance by extraverts on demanding vigilance tasks (Matthews, Davies, & Holley, 1990), although extraversion effects may vary with the nature of the target—non-target discrimination. Extraverts appear to do better with symbolically coded stimuli (e.g., digit sequences) than with physically coded stimuli (e.g., flickering lines).

As might be expected from their superior attentional capacity and STM, extraverts show superior performance to introverts on certain attentionally demanding real-world skills, such as military flying (Bartram & Dale, 1982) and semi-automated mail sorting (Matthews, Jones, & Chamberlain, 1992). Conversely, extraverts may have a higher motor vehicle accident involvement (Furnham, 1992) because the benefits of their processing advantages are outweighted by the disadvantages of recklessness, over-confidence and impulsivity. With regard to intellectual skills, several studies suggest that extraverts are poorer at problem-solving than introverts (see Matthews, 1992), perhaps because they are prone to maladaptive impulsive strategies (Weinman, 1987).

Extraversion as a Noderator of Arousal Effects on Performance

Moderating effects of extraversion on relationships between arousal and performance have been widely reported. The most common finding has been that high arousal is beneficial to the performance of extraverts but detrimental to the performance of introverts (Eysenck & Eysenck, 1985). Eysenck and Eysenck (1985) attribute such findings to chronic under-arousal of extraverts, but research which has assessed the full set of relationships between extraversion, arousal and performance within single studies is inconsistent with this hypothesis (Matthews, 1992). For example, extraversion interacts with self-report arousal in the absence of any negative relationship between the two constructs (e.g., Matthews, Jones, & Chamberlain, 1989). The extraversion x arousal interaction results from the personality-dependence of the relationship between cortical arousal and attentional efficiency. Extraverts tend to show superior performance in states of both high subjective arousal, and low EEG alpha, whereas introverts' performance is maximal in states of subjective fatigue, and high EEG alpha (Matthews et al., 1989; Matthews & Amelang, 1993). A complicating factor is further interaction with time of day. In the evening, the interaction reverses, with extraverts performing better under low arousal, and introverts benefitting from high arousal (Humphreys & Revelle, 1984; Matthews et al., 1989). In fact, in studies which have manipulated time of day, the time of day x extraversion x arousal interaction is so prevalent it may be described as the *modal interaction* (Matthews, 1992).

Not all tasks are sensitive to the modal interaction. Relatively simple tasks requiring routine, somewhat "automatic" encoding of information, such as letter detection and serial reaction, show the effect reliably, but more attentionally demanding, resource-limited tasks do not (Matthews, 1989; Matthews et al., 1989). Harley and Matthews (1992) and Matthews and Harley (1993) present experimental and simulation data which suggest that time of day, extraversion and arousal may interactively affect levels of random noise in a low-level connectionist network for lexical processing. In other words, interaction between extraversion and arousal is primarily associated with individual differences in EIP. Intriguingly, certain intelligence tests are also sensitive to the characteristic interaction (Revelle, Humphreys, Simon, & Gilliland, 1980; Matthews, 1985). In this case, there is evidence that relatively easy "speed" items are more sensitive than more difficult "power" items (Matthews & Dorn, 1995).

Adaptive Benefits of Extraversion and Introversion

At first sight, the extraversion data present the same problem as the intelligence data. The information-processing correlates of extraversion appear to have little in common, with respect to cognitive theory. However, the

patterning of performance associated with extraversion may become more meaningful when considered from an adaptive perspective. Matthews and Dorn (1995) suggest that extraversion is associated with an adaptation towards efficient processing in environments characterised by high rates of information input, especially where information is verbally or symbolically encoded. Interaction with other people typically involves high information flows. In conversation, in addition to the semantic content of verbal utterances, the individual may process other qualities of speech such as inflections, and a variety of non-verbal stimuli. Extraverts' greater competence in processing multiple, unpredictable inputs may enhance social skills and support their greater social interest and sociability (cf. Furnham, 1981, 1992), as well as fitness for demanding jobs such as financial dealing (Kahn & Cooper, 1993) and working on oil rigs (Sutherland & Cooper, 1991), and jobs with a large social component such as sales work (Mount, Barrick, & Strauss, 1994). Self-report data suggest that extraverts are also more sexually active than introverts (Eysenck & Eysenck, 1985), a finding which may, in part, reflect the application of the greater verbal skills of the extravert to impressing and seducing sexual partners. Conversely, introverts appear to be adapted towards functioning efficiently when meaningful stimuli are infrequent, which may facilitate the work activities preferred by introverts such as being a physicist, writer or artist (Cattell & Kline, 1977). There is some evidence that introverts perform better on somewhat monotonous jobs such as card-punch operation (Savage & Stewart, 1972), and that introversion is positively related to creativity in art students (Gotz & Gotz, 1973).

Matthews and Dorn (1995) propose that the overall adaptation of extraverts to high information environments is supported by a number of discrete skills, such as conversational abilities and rapidity of response. Each skill is supported by relevant processing functions known to be associated with extraversion. Superior verbal working memory and availability of attentional resources facilitate conversation skills, whereas adoption of a low response criterion aids rapidity of response. The rather complex time of day x extraversion x arousal interaction may also be adaptive. Matthews and Harley (1993) point out that there is a strong diurnal rhythm in subjective arousal, which is similar in extraverts and introverts (see Thayer, 1989). One of the features of this diurnal variation is that most people are low in arousal for an hour or two in the morning, and for several hours in the evening. Low arousal of this type is somewhat maladaptive, in that it is associated with reduced attentional capacity and poorer performance on demanding attentional tasks (Matthews & Margetts, 1991). One aspect of the modal interaction is that extraverts perform better in the evening when low in arousal. A consequence of this relationship is that extraverts will tend to perform better in the evening than introverts, as demonstrated empirically by Blake (1971), because people are generally in a low arousal state in the evening. From the adaptive perspective, the evening

is the time of day at which parties and other social activities liked by extraverts (Furnham, 1981) take place. Hence, the adaptive function of the modal interaction may be to support extraverts' evening social activities.

In contrast to intelligence, the processing advantages of being an extravert are balanced by disadvantages on certain tasks, such as vigilance and complex problem-solving (Matthews, 1992). Processing functions such as sustaining attention in the absence of stimulation, and the capacity to reflect on problems, may be adaptive in environments which deliver motivationally-significant stimuli infrequently. The introvert may be particularly successful in, for example, artistic, literary and academic occupations in which performance must be directed by some relatively long-term plan, rather than by the task demands imposed by processing of external stimuli. For example, introverts show relatively low levels of educational attainment until age 12, but relatively high levels of attainment after age 15 (Furnham, 1992). It has been proposed that learning in younger children is more of a social activity, whereas at more advanced ages the ability to organise and pursue programmes of individual study becomes important.

Adaptive Relevance of Psychobiological Correlates of Extraversion

Extraversion-introversion may be supported by individual differences in psychobiological as well as cognitive functioning. Introverts appear to be more easily aroused than extraverts, so that they tend to show stress reactions to high levels of stimulation, but are more tolerant of sensory deprivation (Eysenck & Eysenck, 1985; Stelmack, 1990) Similarly, extraverts appear to be more sensitive to reward signals, introverts to punishment signals (Gray, 1981). As argued above, it is unlikely that correlations between extraversion and arousal directly mediate extraversion effects on performance. However, such psychobiological correlates of extraversion may support the overall adaptation, by conferring resistance to the stress liable to be induced by high levels of stimulation. Likewise, introverts may be less susceptible to boredom-induced stress.

Conclusions: Extraversion and Practical Intelligence

The implications of the experimental data for intelligence research is clear. Extraverts show greater adaptive or practical intelligence in certain environments, whereas introverts show more intelligent behavior and processing in other environments. There is an adaptive tradeoff, such that extraversion and introversion are associated with a degree of adaptive specialization. In contemporary society, the "natural" environment for the extravert is the disco or the dealing room, whereas the introvert is equipped for the library or the lighthouse. In contrast, the ambivert is an all-rounder who is moderately well-fitted to a variety of environments. Laboratory studies

of extraversion and performance pick up these adaptations somewhat indirectly, but are sufficiently adequate to indicate the nature of the cognitive functions enabling adaptation.

NEUROTICISM, PERFORMANCE AND ADAPTATION

Neuroticism, Information-Processing and Skill

The second major personality trait of interest is neuroticism, which, within the Big Five psychometric framework, may be seen as equivalent to trait anxiety. Neuroticism tends to be associated with general impairments in performance, although, as with intelligence, scrutiny of the data shows that individual differences are more pronounced on some tasks than others. Matthews and Dorn (1995) suggest that neuroticism/anxiety effects on EIP are fairly weak, and that tasks requiring short-term memory and attentional resources are those most likely to be impaired (see Eysenck, 1992, and Mueller, 1992, for other recent reviews). There is also some evidence that anxious individuals may be impaired in use of executive control processes (e.g., Weinman, Elithorn, & Cooper, 1985), and they sometimes show a more general behavioral passivity (Geen, 1987). Like low intelligence individuals, anxious subjects tend to learn better in more structured environments in which the role of executive processing is reduced (Cronbach & Snow, 1977). Hence, there is some overlap between functions sensitive to neuroticism/anxiety and to psychometric intelligence, although detrimental effects of general anxiety on intelligence test performance are relatively weak (Matthews, 1986). Impairment of complex performance may be more evident under stressful conditions, such as time pressure (Morris & Liebert, 1969) and evaluation (Mueller, 1992). It is often supposed that the worry and intrusive thoughts associated with neuroticism/anxiety divert attentional resources or working memory space away from task performance (Eysenck, 1992; Sarason, Sarason, & Pierce, 1990). Anxiety is also associated with tendencies to appraise the self negatively, and to direct selective attention towards threat stimuli (Wells & Matthews, 1994).

With respect to real-world behavioral criteria, neuroticism tends to be only weakly related to dysfunction. For example, a review by Hough (1992) reports averaged correlations between "adjustment" (low neuroticism) and job proficiency ($r = .09$, $N = 35{,}148$), training success ($r = .12$, $N = 8{,}685$), educational success ($r = .20$, $N = 70{,}588$) and law aiding behavior ($r = .41$, $N = 36{,}210$). The small magnitude of all but the last of these correlations suggests that, in general, high neuroticism is only a minor handicap in educational and occupational settings, though neurotic individuals may be somewhat more prone to delinquency. Neuroticism may somewhat more strongly related to accidents in industrial settings (Hansen, 1989; Sutherland

& Cooper, 1991), and to performance of certain demanding tasks such as military flying (Bartram & Dale, 1982), although the relationship between neuroticism and motor vehicle accident involvement seems inconsistent (Matthews, Dorn, & Glendon, 1991). To some degree, real-world correlates of neuroticism overlap with those of introversion, a finding which may reflect the relationship of both traits to stress vulnerability. However, in the case of introversion, vulnerability may derive primarily from low psychophysiological tolerance of stimulation (Eysenck & Eysenck, 1985) and information-processing limitations (Matthews, 1992), whereas in neurotics it derives from styles of cognitive appraisal and coping (Wells & Matthews, 1994).

Adaptive Significance of Neuroticism

Neuroticism is usually seen simply as a personal deficiency, a view bolstered by its association with mental illness and behavioral problems (Eysenck & Eysenck, 1985). However, in recent years there has been recognition that anxiety and worry may have various adaptive functions, such as anticipating and preparing for threats (Wells, 1994; Wells & Matthews, 1994). Wells (1994) also reviews evidence suggesting that worry, which is strongly correlated with neuroticism, may provide short-term relief from other types of stress symptom such as unpleasant images. Matthews and Dorn (1995) characterize neuroticism as related to an adaptive trade-off between awareness of threat (particularly social threat) and insensitivity to criticism. The neurotic individual is perhaps adapted to environments in which threats are disguised or infrequent, but nevertheless important. For example, in dealing with people from a different culture, the neurotic person may be more careful to avoid causing offence, and more sensitive to cues that a cultural norm is being violated. There is some evidence that neuroticism (together with introversion) is associated with greater sensitivity to the emotions of others, provided the context is non-stressful (Cunningham, 1977; Duckworth, 1975). Conversely, the emotionally stable person will cope better with environments providing overtly negative stimuli, and with stressors such as being evaluated.

The typical laboratory experiment, in which evaluation is implicitly or explicitly part of the context, may be an environment to which the neurotic or anxious individual is not well-suited. Studies of test anxiety show that when efforts are made to reassure performers, anxiety may actually benefit performance (see Sarason et al., 1990). Occasionally, anxiety is associated with superior performance on easy tasks, possibly because anxious subjects are more motivated when success is anticipated (Eysenck, 1982). If, in real life, neuroticism is primarily maladaptive, one might wonder whether contemporary society, with its emphasis on performance evaluation and social competition, is, in general, an environment which favors emotionally stable individuals. In many real-world settings, such as taking examinations, neuroticism may indeed be associated with reduced

adaptive intelligence, but there may also be contexts in which neuroticism is adaptive. There are several studies which show that neuroticism predicts academic success in higher education; McKenzie (1989) reports data suggesting that neuroticism is beneficial in this context if the individual also has effective coping strategies, but not otherwise.

CONCLUSIONS

The data reviewed illustrate the disadvantages of a purely cognitive-theoretical approach to explaining information-processing correlates of individual difference factors. Although application of concepts from cognitive psychology is descriptively rich and informative, this approach fails to provide any general explanation for the particular cognitive patternings (cf. Hockey, 1984) of processing efficiencies associated with intelligence and personality. The adaptive approach suggests that cognitive patternings serve to support skills which contribute to adaptation to specific environments. Figure 1 summarizes some of the relationships tentatively identified in this chapter, although it is emphasised that these are somewhat speculative because of the shortage of direct evidence on the adaptive significance of individual differences in information processing.

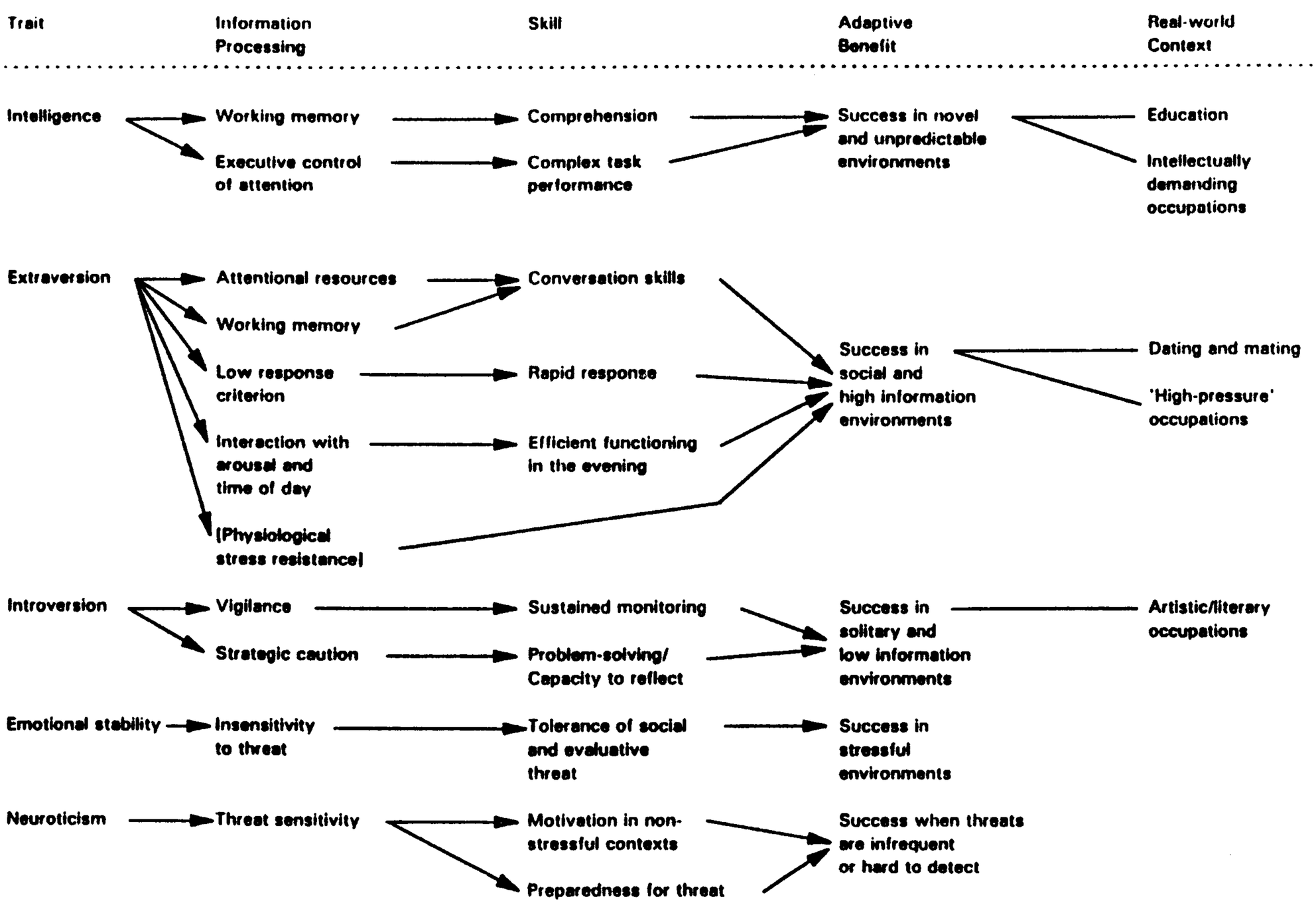


Figure 1. A Schematic Representation of How Information-processing and Skills May Support Adaptive Functions of Individual Difference Traits

In the case of intelligence, superior working memory and executive processing support the acquisition of complex verbal and cognitive skills (e.g., Ackerman, 1988), which in turn support adaptation to novel environments requiring systematised knowledge, such as educational contexts and intellectually demanding occupations. Intelligence may correlate with EIP, not because elementary speed of processing is directly adaptive, but because speed limits capacity to acquire more complex processing routines (Anderson, 1992), and so provides an empirical marker for intelligence. The essential difference between intelligence and personality is that ability rarely relates to poorer performance except under highly-contrived conditions (see Jensen, 1980), whereas extraversion and neuroticism relate to superior performance in some contexts and tasks, but poorer performance in others. For example, extraverts are adapted to stimulating, high-information environments, and possess processing characteristics such as good verbal working memory which support this adaptation. Similarly, introverts may be adapted to low-information environments, and extended processing in the absence of meaningful stimulation. The experimental data suggest that neurotics may be adapted to low social-threat environments, and emotionally stable individuals to high social-threat environments, but evidence is lacking on the real-world contexts in which the two types of individual may be advantaged (and so no entries of this type are made in Figure 1). In general, personality relates to degrees of specialization in adapting to important human contexts, which will be reflected in inter-contextual variation in "practical" intelligence, as defined by Sternberg (1985).

A fuller understanding of individual differences in adaptation requires consideration of the genetic basis for phenotypic variabiliity; there is extensive evidence for a heritable contribution to intelligence, extraversion and neuroticism (Eysenck & Eysenck, 1985). Two evolutionary views of intelligence are possible. One possibility is that the steady temporal increase in brain size and encephalization evident in the fossil data (Stringer, 1984) continues to the present, so that intelligence remains biologically as well as behaviorally adaptive. The alternative is that the performance advantages of intelligence are balanced by disadvantages in other domains, such as reduced fertility (see Brody, 1992) or by other biological costs associated with brain functions related to intelligence (see Eysenck, 1992). Another possible disadvantage is that gaining the benefits of a high-intelligence genotype requires investment of time in knowledge acquisition, which might be better allocated to more directly adaptive behaviors, such as seeking a mate. Future research requires attention both to the real-world contexts in which intelligence facilitates success, and to potential hidden costs of high intelligence.

In the case of personality, the evidence for heritability (Eysenck & Eysenck, 1985) implies that individuals are predisposed by their genetics to particular settings of the adaptive tradeoffs associated with personality. Both cognitive

and psychobiological components of personality may be influenced in this way. At the same time, personality tracks learning and experience to some extent. For example, salespersons become more extraverted as a result of their work (Turnbull, 1976), and neuroticism seems to be as much a symptom as a cause of depression (Barnett & Gotlib, 1988). In other words, genetics provides the individual with a set of basic characteristics which may sometimes by over-ridden by environmental factors, although in the population as a whole phenotype correlates with genotype. Finally, it is important to emphasise that the adaptive approach as outlined here is primarily concerned with behavior rather than with genetics. Even if personality had no genetic basis at all, the individual would still be faced with the adaptive question of how and whether to specialise in dealing with the major contexts for life success, such as social interaction and performance evaluation. The primary task for differential psychologists is to characterise the major adaptive decisions to be made, and to relate decisional choices to psychometric measures of personality and intelligence.

REFERENCES

Ackerman, P.L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117*, 288-318.

Ackerman, P.L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology, 77,* 598-614.

Alderton, D.L., & Larson, G.E. (1994). Cross-task consistency in strategy use and the relationship with intelligence. *Intelligence, 18,* 47-66.

Amelang, M., & Ullwer, U. (1991). Correlations between psychometric measures and psychophysiological as well as experimental variables in studies on extraversion and neuroticism. In J. Strelau & A. Angleitner (Eds.), *Explorations in temperament*. New York: Plenum.

Anderson, M. (1992). *Intelligence and development: A cognitive theory*. Oxford: Blackwell.

Arthur, W., Jr., Barrett, G.V., & Alexander, R.A. (1991). Prediction of vehicular accident involvement: A meta-analysis. *Human Performance*, *4*, 89-105.

Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press.

Barnett, P.A., & Gotlib, I.H. (1988). Psychosocial functioning and depression: Distinguishing among antecedents, concomitants, and consequences. *Psychological Bulletin, 104,* 97-126.

Bartram, D., & Dale, H. (1982). The Eysenck Personality Inventory as a selection test for military pilots. *Journal of Occupational Psychology, 55,* 287-296.

Blake, M.J.F. (1971). Temperament and time of day. In W.P. Colquhoun (Ed.), *Biological rhythms*. London: Academic.

Blanco, M.J., & Alvarez, A.A. (1994). Psychometric intelligence and visual focused attention: Relationships in nonsearch tasks. *Intelligence, 18,* 77-106.

Brand, C. (1994). Open to experience—closed to intelligence: Why the "Big Five" are really the comprehensive six. *European Journal of Personality, 8,* 299-310.

Brand, C., & Deary, I.J. (1982). Intelligence and "inspection time." In H.J. Eysenck (Ed.), *A model for intelligence* (pp. 133-148). New York: Springer-Verlag.

Brebner, J., & Cooper, C. (1985). A proposed unified model of extraversion. In J.T. Spence & C.E. Izard (Eds.), *Motivation, emotion and personality*. Amsterdam: North-Holland.

Brody, N. (1992). *Intelligence* (2nd ed). New York: Academic.

Brookings, J.B. (1990). A confirmatory factor analytic study of time-sharing performance and cognitive abilities. *Intelligence, 14*, 43-59.

Campione, J.C., Brown, A.L., & Ferrara, R.A. (1982). Mental retardation and intelligence. In R.J. Sternberg (Ed.), *Handbook of human intelligence*. Cambridge: Cambridge University Press.

Cantor, J., Engle, R.W., & Hamilton, J. (1991). Short-term memory, working memory, and verbal abilities: How do they relate? *Intelligence, 15*, 229-246.

Cattell, R.B., & Kline, P. (1977). *The scientific analysis of personality and motivation.* London: Academic.

Ceci, S.J. (1990). *On intelligence... more or less: A bio-ecological treatise on intellectual development*. Englewood Cliffs, NJ: Prentice-Hall.

Costa, P.T., Jr., & McCrae, R.R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13,* 653-665.

Cronbach, L.J., & Snow, R.E. (1977).. *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.

Cunningham, M.R. (1977). Personality and the structure of non-verbal communication of emotion. *Journal of Personality, 45,* 564-584.

Czaja, S.J., Joyce, C.B., & Hammond, K. (1989). A comparative study of text editing programs among a sample of older adults. In *Proceedings of the 33rd Annual Meting of the Human Factors Society* (pp. 179-191). Santa Monica, CA: Human Factors Society.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450-466.

Davies, D.R., Jones, D.M., & Taylor, A. (1984). Selective and sustained attention: Individual and group differences. In R. Parasuraman & D.R. Davies (Eds.), *Varieties of attention*. New York: Academic.

Davies, D.R., & Parasuraman, R. (1982). *The psychology of vigilance*. London: Academic Press.

Deary, I.J., & Matthews, G. (1993). Personality traits are alive and well. *The Psychologist, 6,* 299-311.

Dempster, F.N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin, 89*, 63-100.

Duckworth, D.H. (1975). Personality, emotional state and perception of nonverbal communications. *Perceptual and Motor Skills, 40,* 325-326.

Duncan, J. (in press). Attention, intelligence and the frontal lobes. In M. Gazzaniga (Ed.), *The Cognitive Neurosciences*. Cambridge, MA: MIT Press.

Duncan, J., Williams, M.I., Nimmo-Smith, M.I., & Brown, I. (1993). The control of skilled behaviour: Learning, intelligence, and distraction. In D. Meyer & S. Kornblum (Eds.), *Attention and performance XIV*. Cambridge, MA: MIT Press.

Duncan, O.D., Featherman, D.L., & Duncan, B. (1972). *Socioeconomic background and achievement*. New York: Seminar Press.

Egan, D.E., & Gomez, L.M. (1985). Assaying, isolating, and accommodating individual differences in learning a complex skill. In R.F. Dillon (Ed.), *Individual differences in cognition,* (pp. 173-217). New York: Academic.

Egan, V. (1994). Intelligence, inspection time and cognitive strategies. *British Journal of Psychology, 58,* 305-316.

Eysenck, H.J. (1992). Intelligence: The one and the many. In D.K. Detterman (Ed.), *Current topics in human intelligence, Vol. 2: Is mind modular or unitary?* (pp. 83-116). Norwood: Ablex.

Eysenck, H.J., & Eysenck, M.W. (1985). *Personality and individual differences: A natural science approach*. New York: Plenum.

Eysenck, M.W. (1982). *Attention and arousal: Cognition and performance.* London: Springer.

Eysenck, M.W. (1992). *Anxiety: The cognitive perspective.* Hillsdale, NJ: Erlbaum.

Ferretti, R.P., & Cavalier, A.R. (1991). Constraints on the problem solving of persons with mental retardation. *International Review of Research in Mental Retardation, 17*, 153-192.

Fleishman, E.A., & Quaintance, M.K. (1984). *Taxonomies of human performance.* New York: Academic.

Furnham, A. (1981). Personality and activity preference. *British Journal of Social Psychology, 20*, 57-68.

Furnham, A. (1992). *Personality at work: The role of individual differences in the workplace.* London: Routledge.

Gathercole, S.E., & Baddeley, A.D. (1993). *Working memory and language.* Hillsdale, NJ: Erlbaum.

Geen, R.G. (1987). Test anxiety and behavioral avoidance. *Journal of Research in Personality, 21*, 481-488.

Gotz, K., & Gotz, K. (1973). Introversion-extraversion and neuroticism in gifted and ungifted art students. *Perceptual and Motor Skills, 36*, 675-678.

Gray, J.A. (1981). A critique of Eysenck's theory of personality. In H.J. Eysenck (Ed.), *A model for personality.* London: Springer.

Hansen, C. (1989). A causal model of the relationship among accidents, biodata, personality, and cognitive factors. *Journal of Applied Psychology, 74,* 81-90.

Harley, T.A., & Matthews, G. (1992). Interactive effects of extraversion, arousal and time of day on semantic priming: Are they pre-lexical or post-lexical? *Personality and Individual Differences, 13,* 1021-1029.

Hockey, G.R.J. (1984). Varieties of attentional states: The effects of environment. In R. Parasuraman & D.R. Davies (Eds.), *Varieties of attention.* New York: Academic.

Hough, L.M. (1992). The "Big Five" personality variables-construct confusion: description versus prediction. *Human Performance, 5*, 139-155.

Howe, M.J.A. (1989). *Fragments of genius: The strange feats of idiots savants.* London: Routledge.

Humphreys, M.S., & Revelle, W. (1984). Personality, motivation and performance: A theory of the relationship between individual differences and information processing. *Psychological Review, 91*, 153-184.

Hunter, J.E., & Hunter, R.F. (1984). Validity and utility of alternative predictors of a job performance. *Psychological Bulletin, 96*, 72-98.

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

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

Jerison, H.J. (1982). The evolution of biological intelligence. In R.J. Sternberg (Ed.), *Handbook of human intelligence.* Cambridge: Cambridge University Press.

Kahn, H., & Cooper, C.L. (1993). *Stress in the dealing room: High performers under pressure.* London: Routledge.

Koelega, H.S. (1992). Extraversion and vigilance performance: 30 years of inconsistencies. *Psychological Bulletin, 112*, 239-258.

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

Kyllonen, P.C., & Christal, R.E. (1990). Reasoning ability is (little more than) working memory capacity?! *Intelligence, 14*, 389-433.

Larson, G.E., & Alderton, D.L. (1992). The structure and capacity of thought: Some comments on the cognitive underpinnings of *g*. In D.K. Detterman (Ed.), *Current topics in human intelligence, Vol. 2: Is mind modular or unitary?* (pp. 141-146). Norwood: Ablex.

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

Matthews, G. (1985). The effects of extraversion and arousal on intelligence test performance. *British Journal of Psychology, 76*, 479-493.

Matthews, G. (1986). The effects of anxiety on intellectual performance: When and why are they found? *Journal of Research in Personality, 20*, 385-401.

Matthews, G. (1989). Extraversion and levels of control of sustained attention. *Acta Psychologica, 70*, 129-146.

Matthews, G. (1992). Extraversion. In A.P. Smith & D.M. Jones (Eds.), *Handbook of human performance. Vol. 3: State and trait*. London: Academic.

Matthews, G., & Amelang, M. (1993). Extraversion, arousal theory and performance: A study of individual differences in the EEG. *Personality and Individual Differences, 14*, 347-363.

Matthews, G., Davies, D.R., & Holley, P.J. (1990). Extraversion, arousal and visual sustained attention: The role of resource availability. *Personality and Individual Differences, 11*, 1159-1173.

Matthews, G., Davies, D.R., & Holley, P.J. (1993). Cognitive predictors of vigilance. *Human Factors, 35*, 3-24.

Matthews, G., & Dorn, L. (1989). IQ and choice reaction time: An information processing analysis. *Intelligence, 13*, 299-317.

Matthews, G., & Dorn, L. (1995). Personality and intelligence: Cognitive and attentional processes. In D. Saklofske & M.S. Zeidner (Eds.), *International handbook of personality and intelligence*. New York: Plenum.

Matthews, G., Dorn, L., & Glendon, A.I. (1991). Personality correlates of driver stress. *Personality and Individual Differences, 12,* 551-556.

Matthews, G., & Harley, T.A. (1993). Effects of extraversion and self-report arousal on semantic priming: A connectionist approach. *Journal of Personality and Social Psychology, 65*, 735-756

Matthews, G., Jones, D.M., & Chamberlain, A.G. (1989). Interactive effects of extraversion and arousal on attentional task performance: Multiple resources or encoding processes? *Journal of Personality and Social Psychology, 56*, 629-639.

Matthews, G., Jones, D.M., & Chamberlain, A.G. (1992). Predictors of individual differences in mail coding skills, and their variation with ability level. *Journal of Applied Psychology, 77*, 406-418.

Matthews, G., & Margetts, I. (1991). Self-report arousal and divided attention: A study of performance operating characteristics. *Human Performance, 4,* 107-125.

Matthews, G., & Oddy, K. (1993). Recovery of major personality dimensions from trait adjective data. *Personality and Individual Differences, 15*, 419-431.

McKenzie, J. (1989). Neuroticism and academic achievement: The Furneaux Factor. *Personality and Individual Differences, 10,* 509-515.

Morris, L.W., & Liebert, R.M. (1973). Effects of anxiety on timed and untimed intelligence tests: Another look. *Journal of Consulting and Clinical Psychology, 33,* 240-244.

Mount, M.K., Barrick, M.R., & Strauss, J.P. (1994). Validity of observer ratings of the Big Five personality factors. *Journal of Applied Psychology, 79,* 272-280.

Mueller, J.H. (1992). Anxiety and performance. In Smith, A.P. & Jones, D.M. (Eds.), *Handbook of human performance. Vol. 3: Trait and state*. London: Academic.

Necka, E. (1992). Cognitive analysis of intelligence: The significance of working memory processes. *Personality and Individual Differences, 13,* 1031-1046.

Nettelbeck, T. (1990). Intelligence does exist. *The Psychologist, 3,* 494-497.

Nettelbeck, T., & Rabbitt, P.M.A. (1992). Aging, cognitive performance, and mental speed. *Intelligence, 16*, 189-205.

Neubauer, A.C. (1991). Intelligence and RT: A modified Hick paradigm and a new RT paradigm. *Intelligence, 15*, 175-192.

Norman, D.A., & Bobrow, D.B. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7,* 44-64.

Rehberg, R.A., & Rosenthal, E.R. (1978). *Class and merit in the American high school: An assessment of the revisionist and meritocratic arguments.* New York: Longman.

Revelle, W., Humphreys, M.S., Simon, L., & Gilliland, K. (1980). The interactive effect of personality, time of day and caffeine: A test of the arousal model. *Journal of Experimental Psychology: General, 109,* 1-31.

Richards, G. (1987). *Human evolution: An introduction for the behavioural sciences.* London: Routledge.

Roberts, R.D., Beh, H.C., & Stankov, L. (1988). Hick's law, competing task performance, and intelligence. *Intelligence, 12,* 101-120.

Saklofske, D.H., & Kostura, D.D. (1990). Extraversion-introversion and intelligence. *Personality and Individual Differences, 11,* 547-552.

Sarason, I.G., Sarason, B.R., & Pierce, G.R. (1990). Anxiety, cognitive interference, and performance. *Journal of Social Behaviour and Personality, 5,* 1-18.

Savage, R., & Stewart, R. (1972). Personality and the success of card-punch operators in training. *British Journal of Psychology, 63,* 445-450.

Shute, V.J., Glaser, R., & Raghaven, K. (1989). Inference and discovery in an exploratory laboratory. In P.L. Ackerman, R.J. Sternberg & R. Glaser (Eds.), *Learning and individual differences: Advances in theory and research* (pp. 279-326). New York: Freeman.

Snow, R.E. (1989). Aptitude-treatment interaction as a framework for research on individual differences in learning. In P.L. Ackerman, R.J. Sternberg & R. Glaser (Eds.), *Learning and individual differences: Advances in theory and research* (pp. 13-60). New York: Freeman.

Snow, R.E., & Yalow, E. (1982). Education and intelligence. In R.J. Sternberg (Ed.), *Handbook of human intelligence.* Cambridge: Cambridge University Press.

Spearman, C. (1923). *The nature of "intelligence" and the principles of cognition.* London: MacMillan.

Spilsbury, G.A. (1992). Complexity as a reflection of the dimensionality of a task. *Intelligence, 16,* 31-45.

Stankov, L. (1989). Attentional resources and intelligence: A disappearing link. *Personality and Individual Differences, 10,* 957-968.

Stelmack, R.M. (1990). Biological bases of extraversion: Psychophysiology evidence. *Journal of Personality, 58,* 293-311.

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

Sternberg, R.J., Conway, B.E., Ketron, J.L., & Bernstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology, 41,* 37-55.

Sternberg, R.J., & Wagner, R.K. (1989). Individual differences in practical knowledge and its application. In P.L. Ackerman, R.J. Sternberg & R. Glaser (Eds.), *Learning and individual differences: Advances in theory and research* (pp. 255-278). New York: Freeman.

Stringer, C. (1984). Human evolution and biological adaptation in the Pleistocene. In R. Foley (Ed.), *Hominid evolution and community ecology: Prehistoric human adaptation in biological perspective,* (pp. 55-84). London: Academic.

Sutherland, V.J., & Cooper, C.L. (1991). Personality, stress and accident involvement in the offshore oil and gas industry. *Personality and Individual Differences, 12,* 195-204.

Swanson, L.H. (1993). Executive processing in learning-disabled readers. *Intelligence, 17,* 117-150.

Thayer, R.E. (1989). *The biopsychology of mood and arousal.* Oxford: Oxford University Press.

Turnbull, A. (1976). Selling and the salesman: Prediction of success and personality change. *Psychological Reports, 38,* 1175-1180.

Vicente, K.J., & Williges, R.C. (1988). Accommodating individual differences in searching a hierarchical file system. *International Journal of Man-Machine Studies, 29,* 647-668.

Weinman, J. (1987). Non-cognitive determinants of perceptual problem-solving strategies. *Personality and Individual Differences, 8*, 53-58.

Weinman, J., Elithorn, A., & Cooper, R. (1985). Personality and problem solving: The nature of individual differences in planning, scanning and verification. *Personality and Individual Differences, 6*, 453-460.

Wells, A. (1994). Attention and the control of worry. In G.C.L. Davey & F. Tallis (Eds.), *Worrying: Persepctives on theory, assessment and treatment.* Chichester: Wiley.

Wells, A., & Matthews, G. (1994). *Attention and emotion: A clinical approach.* Hillsdale, NJ: Erlbaum.

Westerman, S.J., Davies, D.R., Glendon, A.I., Stammers, R.B., & Matthews, G. (in preparation). Individual differences in skilled word processing.

Yee, P.L., Laden, B., & Hunt, E. (1994) The coordination of compensatory tracking and anticipatory timing tasks. *Intelligence, 18,* 259-288.

THE CONCEPT OF "INTELLIGENCE" WITHIN NEUROPSYCHOLOGY:

COMMENTS AND REFLECTIONS

Maria Levander

INTRODUCTION

The present chapter aims at bringing out certain aspects of the concept of intelligence within the clinical and research domain of adult neuropsychology. Many issues are relevant for both clinical and research work but will, for clarity of presentation, be treated under separate headings.

Because of the multitude of intellectual functions there is no general intellectual function that can be denoted intelligence. IQ scores which are composite scores, and derived from intelligence tests, are therefore not useful in describing the intellectual performance of normals or of patients with neuropsychological disorders. Accordingly, the term intelligence will not be used in the following. Instead it will be elaborated upon in terms of cognitive functioning which is based on the processing of numerous modules or microprocessors relatively independently localized within the brain (cf. Mountcastle, 1978 cited in Gardner, 1983). Cognitive functioning may thus

Advances in Cognition and Educational Practice, Volume 4, pages 201-219.

ISBN: 0-7623-0105-8

be perceived in terms of an orchestrated interdependent activity of multiple cognitive processors or modules, each responsible for different functions that are sometimes overlapping one another (see Allen, 1983). As yet, we do not know the number and content of all these modules engaged in human cognitive functioning. Furthermore, it is questionable whether all abilities readily lend themselves to this type of fractionation. As pointed out by Eysenck and Keane (1990) the modules do perhaps not "really" exist. Nevertheless they may be regarded as convenient devices so as to clarify our understanding of cognitive functioning.

THEORETICAL AND METHODOLOGICAL ISSUES IN A CLINICAL CONTEXT

Assessment in neuropsychology can be carried out for different reasons. Hence, it can be for a diagnostic purpose or for assessing present cognitive impairments or for change if it is to be monitored as a result of a rehabilitation program, or for research purposes.

In the context of clinical neuropsychological assessment the most common and obvious way to assess cognitive functioning level and cognitive profile, possibly indicating cognitive dysfunctions is by means of formal psychological tests of which the Wechsler Adult Intelligence Scale (WAIS) is most often used. Despite certain modifications of the test battery, it is still a crude instrument.

IQ measures alone do not adequately reflect the complexity of preserved cognitive capacity and impairments that may have emerged after for example, closed head injury (cf. Newcombe, 1987). However, WAIS provides the examiner with hypotheses for subsequent assessment. With the construction of the neuropsychological version of the WAIS battery that presents a process approach to cognitive functions (WAIS-RNI; Kaplan, Fein, Morris, & Delis, 1991) some progress has, however, been made since the results are not only treated quantitatively.

At each assessment it is relevant to ask what type of cognitive activity is being assessed. Whyte (cited in Gordon, 1990) distinguishes between six different cognitive activities ranging from the most basic form of "cognitive operation" followed by "cognitive processes," "single skills," "meta skills," "automatic performance" to the most complex "global function." It is the activity of the cognitive processes that are being assessed by the neuropsychological tests.

It is characteristic of human performance that is, the output of cognitive processing, that a variety of different strategies can be employed in most tasks (e.g., Levander, Tegnér, & Caneman, 1993). However, typically paper and pencil tests provide measures of performance reflecting skills (ability) but rarely allow detailed analyses of problem solving strategies (e.g., speed vs. accuracy).

The absence of quantitative strategy parameters for these tests precludes an analyses of their relative contribution to test performance. Performance is not equal to competence level (capacity) however, the latter only being indirectly observable. Competence is better considered in terms of a structural model describing the rules whereby it functions whereas performance refers to how it functions in practice (Chomsky, cited in Cohen, 1983). Computerization of certain neuropsychological tests appears to make possible a more detailed analysis of performance in terms of skill, task solving strategies and laterality aspects (cf. Sternberg, 1977) the latter as inferred from hemispheric activation or priming (Kinsbourne, 1970; see also Levander, 1987; Levander, Levander, & Schalling, 1989; Levander & Levander, 1990).

Since available standardized and objective tests do not readily lend themselves to repeated measures computerized tests may also be a more appropriate tool for diagnosing and monitoring cognitive changes as a consequence of pharmacological interventions although it can be problematic to find an entirely appropriate dependent measure (cf. Cope, 1990). It may be suggested that this dependent measure can be established at different levels in the hierarchy of cognitive activity that is, with computerization the patient can practice on the same test but with different complexity thus tapping another aspect or level of cognitive functioning. There are other problems associated with the measurement of cognitive change. As pointed out by Miller (1992) it is necessary to find out the particular functions that have changed, that is, declined or improved, and the degree of change that has occurred. The computerized tests need therefore to have good construct validity. This is crucial since cognitive activity consists of a number of different functions and it is most likely that pharmacological interventions will have different effects on the constituting modules or processors whichever they may be. The use of computers can, however, at times cause problems for physically disabled patients, thus rendering assessment or training difficult. Most response devices can, however, easily be modified to suit subjects with various motor disabilities.

Clinical Application

The clinical application of the neuropsychological tests concern sometimes primarily differential diagnosing or merely evaluating the presence of cognitive deficits in a patient (see under assessment). Thus the examiner expects to find cognitive or behavioral alterations in a specific patient. With this expectation in mind the examiner runs the risk of overlooking the fact that manifested behavioral and/or cognitive alterations may be the patient's affective reaction to the trauma or the disease rather than being a consequence of brain damage (cf. Benton, 1987). One type of differential diagnosis refers to that of hemispheric localization of the epileptic focus/foci by means of neuropsychological tests, when surgery has been considered as the appropriate tool for

alleviating epilepsy. Unfortunately there are, as yet, no neuropsychological instruments that can detect foci in a valid way. Although the results from different radiological measurements are not always congruent, it is perhaps then redundant to provide neuropsychologiocal evidence of a possible localization of seizure focus if several of the radiological examinations point to a specific area. On the other hand by assisting the medical personel in trying to establish the accurate hemispheric side and site of the focus/foci the neuropsychologist can convey information that is presumably of complementary value in terms of criterion validity.

At other times the neuropsychologist is confronted with the problem of deciding whether detection and quantification of impairment significantly deviates from the premorbid level in order to arrive at recommendations for rehabilitation of specific dysfunctions. This is problematic in the sense that there is a substantial variation in cognitive capacity in the general population and comparing an individuals test score with the relevant test norm is only of limited value. For example, an individual with above average premorbid intellectual capacity may achieve within the normal range and yet be unable to continue his current work. There are principally two methods of estimating a patient's expected or premorbid level of performance. The most common method is to use present ability measures which is based on the two following assumptions: an individual's performance on one cognitive measure will permit some estimation of another; and cerebral dysfunction affects some tests of cognitive functioning considerably more than others. The other complementary method uses a test that has been specifically designed to estimate premorbid ability. Often this is a reading or a vocabulary test. However, in the present author's experience the estimation of premorbid level in patients with traumatic brain injuries has proven problematic since many of the patients sustain damage to verbal functions. History data from the patient and/or from family, if there is one, can be of great value even though it does not provide sufficient information. In any case informal preinjury/premorbid information is valuable in individual cases, but not recommended for comparisons between groups, one reason being that there is too low a correlation between such information and cognitive status (Dikmen & Temkin, 1987). A more formal way of obtaining preinjury/premorbid information is to examine test scores from the army, a procedure particularly relevant for males since they are more often within the army and are more frequently afflicted by traumatic head injuries than are females.

There are other problems of considerable neuropsychological significance pertaining specifically to patients with traumatic head injuries. Most often the tests are introduced in a structured setting rendering assessments of executive functions such as planning, alertness, self-correcting behavior almost impossible. In addition subtler problems associated with brain damage such as fatigue, drowsiness, attentional deficits and so on often go undetected

although having a general effect of reducing the ability to cope with tests, particularly those that are difficult. These aforementioned factors may indicate a development of compensatory strategies which obscure the direct effects of brain damage on the cognitive system. It is also the author's experience that alcohol abuse, particularly among the middle aged, irrespective of sex, aggravates cognitive impairments after head injury. Not only being the cause of head trauma in certain cases, the abuse of alcohol does thus have an additive effect on cognitive dysfunctioning making a differentiation between cognitive dysfunctions due to the trauma and to alcohol respectively extremely difficult. The problem resembles that seen in the evaluation of epilepsy surgery that is, although a generally lower dose after surgery, the anticonvulsants conceal the effect of the surgery per se on cognitive functioning. One way of elucidating the effect of surgery on cognitive capacity would, of course, be to compare cognitive performance prior to and after surgery during a period when the patient is taking the same kind and the same dose of medication. The comparison is, however, difficult to make since, after surgery, the dose will hardly ever reach that taken just before surgery which is usually when the neuropsychologist sees the patient. When on lower doses the decision about surgery may never even be relevant in the first place, so a visit to the neuropsychologist never becomes necessary. This contrasts with child neuropsychologists who see their patients on a more regular basis during a longer period of time.

Concerning the compensatory strategies mentioned above, investigators should differentiate specific from generalized neuropsychological impairments (Strauss & Allred, 1987) by looking for double dissociations across different cognitive tests (Eysenck & Keane, 1990). A crucial dissociation occurs when a patient performs normally on one test and at an impaired level at another one. Different possible interpretations then become possible, of which one is the presence of an unusual premorbid cognitive system in a certain patient. If the same dissociation is shown in several patients, then it is, however, more likely that current test performance is a result of brain damage and not of a similar compensatory mechanisms or unusual cognitive make-up in all patients prior to damage. A double dissociation between two tests, 1 and 2, is shown when one patient performs normally on test 1 and dysfunctionally on test 2, and another patient performs in the opposite manner. Should a double dissociation be present the results can not be explained in terms of differences in difficulty between the tests. An alternative approach to the double dissociation technique is to consider patients demonstrating a tendency to exhibit impaired performance on the same set of tests. If such cognitive dysfunctions across tests are shown, there may be an impairment in some cognitive mechanism underlying performance in all tests. It may also suggest that, anatomical structures localized close to each other and harboring quite different cognitive processes necessary for solving a certain testproblem, have

been damaged so that impairment to one cognitive process is accompanied by dysfunction to all of the others. The problem then is how to decide whether a manifested dysfunctional performance is a consequence of damage to adjacent anatomical structures or impairment of some common cognitive process. According to Ellis and Young (1988) an appropriate way of finding out if an association has occurred because of anatomical closeness in cognitive mechanisms is to find a patient who does not show an impairment across tests on which most patients are performing at a low level.

Rehabilitation

The initial task in cognitive rehabilitation is to identify which are the problems that need to be treated. This procedure is often cumbersome and difficult, especially in traumatic brain injuries since the head injury itself gives rise to an extensive variation of neurological and medical symtoms, as well as a considerable variation of emotional, vocational, financial and psychosocial factors. When seeking to identify the deficits the cognitive assessment procedure ought to be directed towards the determination of the status of particular functions, and not on syndromes. Otherwise, the assessment may result in an oversimplification based on the doubtful assumption that manifested cognitive dysfunctions reliably imply the status of other functions. The results from the assessment procedures, which should be performed by a neuropsychologist, subsequently need to be communicated to other professionals. All too often the reports are written in a complicated style, often rife with unexplained terminology, thus making it difficult for the reader to extract the essentials. Simplification and condensation is therefore clearly needed.

Once the cognitive problems have been teased out and hopefully made quantifiable the expected outcome of a particular treatment needs to be conceptualized in behavioral terms. Identification and conceptualization is important because the validity of a rehabiliation program may otherwise be questionable. More specifically, if it is not obvious which specific aspect of the treatment program has had an effect or otherwise on the present cognitive dysfunctions, it is impossible to advocate the utility of a certain program for a specific type of cognitive dysfunctions. Moreover, with the purpose of increasing motivation, the patient and his/her family, should, if possible, be engaged in this procedure.

Questions referring to ecological validity are often raised concerning remediation i.e. what is the benefit from cognitive or behavioral training on everyday life? When participating in a program the purpose is to provide the patient with an underlying principle, or knowledge about how to solve problems, which can be generalized to other environments. Stated differently, training should not be maintained to improve test scores but to ameliorate the impaired cognitive processes that underlie dysfunctional behavior. It is, however, unclear

what the specific principle may be in a specific outcome. But by practising less complex cognitive activities (see under assessment), improvement in the more basic functions will most certainly have an impact on a range of skills to which these different basic functions are related (Gordon, 1990).

Evaluation of Treatment

In clinical settings, the administration of several single-case studies (Eysenck & Keane, 1990) or the conducting of small quasi-experimental group studies (Cook & Campell, 1979) with good pre-treatment and post-treatment assessments seem to be a powerful alternative to classical group studies (von Cramon & Matthes-von Cramon, 1990) since it is not acceptable for ethical reasons to assign patients randomly to treatment groups or to have control groups. Eysenck and Keane (1990, p. 505) propose the use of single-case studies when "fairly detailed theories have been constructed." Mateer and Ruff (1990) also consider single-case methodology as a valuable tool regarding evaluation of rehabilitation although they suggest that more meaningful group comparisons are needed since the "acceptance of behavioral rehabilitation may be weakened if the field is based exclusively on single-case studies." It is here suggested that particularly long term follow-up studies may profit from the use of demographically similar control cases in evaluating cognitive/behavioral remediation programs (e.g., Herson & Barlow, 1976; Kazdin, 1980). Thus traumatic brain injuries can be compared with spinal cord injuries since both groups are members of an at-risk population (see under research section in this chapter). Follow-up studies are unfortunately cumbersome to pursue since they are associated with several practical problems such as maintaining continuity, reducing drop-out rate, identifying and specifying what happens between testing (Brooks, 1987).

A few words of caution are here called for. When in a clinical setting patients sometimes seem to be ignorant about whether they are being treated or reassessed after a short period of for example, retreat. This seems often to be the case when using computers in remediation programs usually administered by professionals other than neuropsychologists. Repeated testing must be separated from treatment otherwise prior experience of test and setting, and change of solving strategies can be confounded with treatment effects. In other words, in order to avoid these practice effects clincians need to demonstrate that relearning has an effect on some variable that is above and beyond what might have been expected based on spontaneous recovery. Needless to say this is crucial both from an economic as well as from a motivational point of view. In sum, although focussing on some of the problems associated with cognitive functions within clinical neuropsychology, the above nevertheless clearly shows the necessity of treating the phenomenon of "intelligence" in terms of cognitive functioning instead of intelligence as constituting a unitary concept.

THEORETICAL AND METHODOLOGICAL ISSUES IN A RESEARCH CONTEXT

Recent research and theory are making the complexities of human cognitive functioning increasingly apparent. Consequently, it has become difficult to make theoretically acceptable distinctions between different functions involved in cognitive capacity. It can be argued that different functions are continuously interacting with each other although to a varying extent depending on individual differences in cerebral organization and activation. This suggests that several functions traditionally subordinated in other domains within neuropsychology should be more and better integrated if neurocognitive functioning is to be fully understood. Of relevance would then be e.g. functions pertaining to aspects of sensation, perception and memory. In pursuing research with regard to cognitive functioning several subject characteristics such as handedness and sex probably modify a clearcut relationship between neurocognitive behavior and brain functioning. These factors need therefore to be accounted for by any theory that wishes to explain cognitive behavior and its underlying neural mechanisms. With the above in mind some of the author's studies from pathological and normal cases will hereby be reported and/or commented on.

Patients

In view of the interacting aspects of virtually all cognitive processes and structures, as mentioned above, the following will draw attention to some anatomical structures and neuropsychological functions that might be of neuropsychological cognitive significance and interest in the present context. Accordingly, focus will be on patients with possible impairments to the corpus callosum as a result of neurosurgical intervention or as a result of head trauma; suffering from neglect syndrom due to stroke, and with dysmnestic olfactory functioning as a consequence of temporal lobe epilepsy.

Corpus Callosum and its Possible Cognitive Implications

In neuropsychology the use of brain damaged patients is a common tool for inferring both aspects of normal cognitive functions and the neural basis for such functions. By studying impaired and spared abilities in these patients it is assumed that the function of a particular brain area is established. However, this approach inferred from behavioral dysfunctions, is associated with several shortcomings.

One major problem pertains to the unlikelihood of destruction of a single cognitive system after a lesion without impinging on other systems or disrupting the connections between them (see clinical application). This disruption can

occur within as well as between the hemispheres. The interhemispheric fibers primarily function to join the two hemispheres and include several commissures of which the corpus callosum provides the major connection of neocortical areas. Surprisingly few neuropsychological studies have been made on patients with impaired callosal nerve fibers. The major hurdle is probably that affected patients are not readily evaluated since they often are severly damaged. Studies on individuals with callosal agenesis is however not entirely comparable since in these individuals the anterior commissure probably is greatly enlarged thus compensating for complete or partial absence of the corpus callosum. In addition, individuals born without a corpus callosum might have an unusual cerebral organization with regard to pathways of the nerve fibers and/or of speech. The effects of disruption of callosal fibers can nevertheless be investigated in patients subjected to partial sectioning of the corpus callosum for evacuation of brain tumors (e.g., Levin et al., 1993), or possibly in moderately severed head trauma patients (Levander, in preparation). Specific tests are however needed otherwise the deficits remain unnoticed.

In the former study, in comparison with normal controls, the patients, after evacuation of the tumor, demonstrated subtle motor and/or sensory manifestations of hemispheric disconnection. These subtly manifested interhemispheric dysfunctions tap cognitive functioning to the extent that they might be considered, when not impaired, as necessary prerequisites for optimal cognitive processing in whatever way that might be.

Data, three and six months postinjury, from the moderately head injured patients indicate similar dysfunctioning in transference of motor and sensory information across the corpus callosum when compared with their age-matched controls (normals and spinal cord injured patients). The results also suggest impairment in interhemispheric transfer of language information. Moreover, slowness in motor performance and a much slower cognitive processing of information was revealed among the head injured patients in comparison to the controls, findings that have repeatedly been found. The subtle dysfunctions improved, however, over a three months period in parallel with recovery of cognitive functions and social interactions.

With the purpose of being able to initiate an early and appropriate remediation program it would perhaps be of interest to systematically investigate whether presence of early dysfunctional interhemispheric transference of information as a result of head trauma would be of any descriptive and prognostic value in terms of cognitive capacity, particularly speech, and social interactions. However, it has to be borne in mind that it is quite possible that the mentioned interhemispheric dysfunctions may be attributed to associated damage, and not exclusively to damage of the corpus callosum since callosal lesions are rarely isolated. One way of solving this problem would be to obtain a sufficient variety of cases with moderately severe head injury in order to be able to convincingly attribute similar disconnection

symptoms to common anatomic aspects. Accordingly, future studies will hopefully demonstrate whether the concept of hemispheric disconnection is relevant when dealing with head injury and, perhaps, rehabilitation.

Neglect as Determinant of Cognitive Impairment?

Studying patients with brain damage poses another type of problem, namely that the effects of lesions show considerable variability between patients. The cognitive effects will hence vary with for example, etiology, site and severity of the lesion, and time from onset of the damage or disease. For example, an extremely wide range of cognitive deficits of varying severity can thus be observed following stroke, in itself considered as a wide-ranging vascular disease. Different kinds of perceptual deficits, of which one is visual neglect, can be observed following stroke. It can however again (see earlier) be argued whether perceptual dysfunctions are "really" of cognitive character. At least perception might be considered to underly optimal cognitive capacity since it constitutes the end-product of those inferences that are being made about stimuli presented to our senses.

In visual neglect or inattention, the patient behaves as if the left side of visual space did not exist. This symptom is most common in patients with a posterior right hemisphere stroke, the symptom being particularly severe in the acute stage. For unknown reasons not all patients affected by stroke in this area manifest visual neglect, but for those affected, neglect may persist for a longer period of time.

For example reexamination 1-4 years after stroke showed in the present study, that 14 out of 16 patients with moderate to severe neglect in the acute stage still demonstrated signs of neglect, with one patient manifesting severe neglect that interfered with daily activities (Hjaltason, Tegnér, Tham, & Levander, 1994).

When daily activities have become problematic as a consequence of neglect there are reasons to believe that, especially among the elderly, simple mechanisms (cognitive operations?) underlying cognitive functioning have become impaired. These mechanisms need to be delineated and conceptualized if the nature of neglect and its cognitive implications are to be understood, and if training is going to be successful for these types of patients.

The neglect syndrome can extend to and affect reading. Neglect dyslexia refers to a reading disorder where either part of the text on the left side of the page is omitted, or where the initial letters within words are omitted or misread (Ellis, Flude, & Young, 1987). When reading single words the length of the words are typically preserved in this disorder, that is, patients tend to substitutes the initial letters rather than omitting them so that for example, BEAT is read as HEAT rather than EAT. The explanation for this impaired reading behavior remains unsolved. Tegnér and Levander (1993) demonstrated

that when reading single words, two stroke patients with neglect dyslexia showed preserved word length as predicted. They also read the words as being longer than the target words when there was a change in context, and as shorter when the change was lexical. The findings suggested that any attempt of explaining word length preservation in neglect dyslexia must consider the influences of the contextual and lexical parameters on single word reading.

One important issue that needs to be investigated in the future pertains to the relation and the dissociability between neglect dyslexia and other types of visual neglect (cf. Ellis & Young, 1988). Once the underlying mechanisms in this acquired reading disorder have been delineated and explained it is tempting to apply these explanations to certain issues regarding for example, developmental dyslexia. Such comparisons are however doubtful since developmental dyslexia most likely is associated, to a larger degree, with broader cognitive deficits than is acquired dyslexia despite several similarities in reading behavior. In the different dyslexias, word recognition is also considered from different theoretical perspectives. That is, the acquired neglect dyslexia is explained in terms of, skilled, adult word recognition whereas developmental dyslexia is more appropriately explained within a framework for development of normal reading behavior (Ellis, 1993).

Olfaction and Cognition—of Interest to the Neuropsychologist?

In addition to the opinion that olfaction is too primitive a function to be of any neuropsychological significance, the reluctance to studying olfaction among neuropsychologists probably stems from the fact that humans do not have access to a well articulated representational system for labelling and classifying odors in terms of their perceptual properties thus rendering investigation very difficult. This should not however hamper researchers from investigating olfaction since this modality is interacting, not only with emotion, but with memory and cognitive functions (Richardson & Zucco, 1989).

For the neuropsychologist interested in olfaction it can be difficult to recruit patients suitable for this type of research since olfactory dysfunctions are not that prominent in all patients suffering from neuropsychological deficits. Nonetheless, patients with Korsakoff's psychosis and patients with damage to the temporal lobes have shown to be appropriate since they often suffer from both olfactory and mnestic dysfunctions (e.g., Eskenazi, Cain, Novelly, & Friend, 1983).

From a cognitive point of view one of the most theoretically interesting issues pertains to the notion of olfactory representation and its neural substrate. Though no coherent conclusion is, as yet, established neuroanatomical regions within the right "nonlinguistic" hemisphere appear to be dominant in the identification and retention of odors. This would hence perhaps explain the difficulty whereby odors are labelled and subsequently retrieved.

There is an ongoing study, examining the effect of side of temporal lobe epilepsy on immediate and delayed recognition of both nameable and non-nameable "ecologically natural" odors, and comparing odor recognition before and after temporal lobe resection (Levander, in preparation). The study aims additionally to examine the emotional impact on recognition of odors, and to document to what extent memory for odors is related to cognitive performance.

For the future, this area of research opens up possibilities for investigations, not only of olfaction and its relation to different aspects of cognitive functioning, but also of the neuropsychological significance of olfactory functioning across subjects differing with regard to some of the most relevant factors that is, age, sex, menstruation cycle, smoking behavior and brain damage (cf. Richardson & Zucco, 1989).

Normals

At the beginning of this chapter it was underscored that, in neuropsychology, the concept of intelligence is more appropriately treated in terms of cognitive functioning regardless of whether patients or normals are being examined. From having dealt with patients above, the section as follows deals with cognitive processes in terms of skill, solving strategy and laterality aspects, underlying cognitive functioning, in a study of normals.

A large study with the purpose of exploring neuropsychological factors, particularly in cognitive functioning in a group of left-handed males and females and their right-handed age-matched controls, was conducted some years ago by the author (Levander, 1988). The 102 left-handers and the 102 right-handers were recruited from 921 individuals (aged 17 to 20 years) attending junior college north of Stockholm.

The cognitive neuropsychological tests consisted of verbal (academic skills—comprehension, vocabulary, and spelling; verbal fluency—fluency of writing and of speech, and reading speed) and perceptuomotor tests (tracing and trail making). Three computerized tests (Reaction Time Test, Perceptual Maze Test—visual spatial functioning, Trail Making Test) were employed so as to make possible an analysis, not only of performance, but of solving strategies and side preferences during the solution process (Levander, 1987).

The study also comprised collection of data concerning different aspects of psychophysiology, psychoendocrinology and personality characteristics which are now being analyzed. A follow up study is shortly to take place.

The Concept of Handedness

The use of handedness as an independent factor in cognitive studies derives from the assumption that it reflects cerebral functional lateralization

underlying cognitive functions. Left-handers differ, as a group from right-handers both in degree and in direction of hemispheric asymmetry (for a review see Annett, 1985).

From a phenomenological viewpoint the assessment of handedness would seem easy. However, handedness is an elusive characteristic and in the absence of any valid criterion for what constitutes "handedness" the diversity of methods used for assessing the incidence of left-handedness is a major problem when comparisons between different studies are to be made. This may explain the differences in proportions of left-handedness found in different populations and among males and females. Proportions of left-handedness ranges from 1 to 30 percent with a mode of about 10 percent. In a majority of studies there is a higher frequency of left-handedness in males. In the population under study 4 percent of the respondents reported a consistent use of the left hand in fine motor activities. Males and females did not differ in reported hand preference for unimanual fine motor activities (Levander & Schalling, 1988).

Handedness in Relation to Cognitive Performance and Strategy

It has been claimed that among left-handers there is a greater incidence of low intellectual capacity, poor reading skill and mental retardation (Satz, 1972). Although there is an overrepresentation of left-handers among mentally retarded individuals this is not to say that a high proportion of left-handers are mentally retarded. The majority of left-handers are neurologically and cognitively normal (e.g., Swanson, Kinsbourne, & Horn, 1980). It was, however, argued that despite the absence of gross cognitive differences there may well be important variations in problem solving strategies between the handedness groups. The argument was based on the idea that solving strategies might be connected to cerebral organization (see Allen, 1983).

The results from the study indicated that left-handedness alone (Levander, Levander, & Schalling, 1989) or in interaction with sex (Schalling & Levander, unpublished manuscript) was associated with superior performance in verbal functions. In addition, left-handers outperformed right-handers in visuo-spatial skills and in perceptuomotor functions (Levander, Levander, & Schalling, 1989). This superiority among left-handers may be ascribed to a hemispheric bilateralization of functions that is being of benefit for which the reason is still unknown, among students with above average intellectual capacity (Levander, Levander, & Schalling, 1989; cf. Harshman, Hampson, & Berenbaum, 1983).

When the focus was on the strategy aspect of problem solving there were only minor differences between left-and right-handers.

Possible Moderating Factors of Handedness on Cognitive Performance-Familial Sinistrality and Strength of Left-Handedness

Familial sinistrality (FS) is one of the most intriguing among factors associated with individual differences in cerebral organization of linguistic and visuospatial functions (Hécaen, 1984). Why should the presence of left-handers in the family predict cerebral functional organization in an individual, regardless of his own handedness? The interpretation offered is that presence of left-handers in the family (FS+) is strongly associated with a bilateralization of cerebral language processing. This hypothesis was originally based on clinical aphasia studies showing, among FS+ subjects, the presence of dysphasia regardless of side of lesion and a faster recovery. The concept of bilateralization as it is used in the research literature, is however ambiguous.

In order to elucidate the implication of familial sinistrality for cognitive performance each handedness group in the study was subdivided according to sex and familial sinistrality (FS+=at least one left-handed member in the immediate family; FS-=no left-handers in the immediate family).
Three factor ANOVA's with handedness, FS and sex as independent variables were applied to cognitive data.

The main finding was the consistent superiority of male FS+ subjects in verbal tasks of different kinds, both in verbal fluency and reading speed, and in academic skill measures, as compared to FS- males irrespective of handedness. It is however still not obvious why a bihemispheric linguistic competence should be associated with greater verbal skill. One reason why bilateralization of language functions would result in higher verbal performance might be that verbal tasks require multiple cerebral functions with involvement of right as well as left hemisphere type of processing (Schalling & Levander, unpublished manuscript).

As mentioned above there were only minor differences in strategy between the handedness groups. This result was considered noteworthy in view of the empirical findings suggesting variations in cerebral organization between left- and right-handers. However, left-handers are more heterogeneous in various aspects of possible relevance for cortical functioning which may imply that any consistent relations between cerebral organization and strategy choice may be obscured. With the intent of getting more homogeneous subgroups in order to discern consistent patterns in terms of solving strategies and choice of side preference (laterality aspect) in the data the left-handed males and females were thus subdivided according to familial sinistrality and strength of left-handedness (weak/inconsistent or strong/consistent; Levander & Levander, 1990).

It was only in the rather complex maze test that effects of strategy and side preferences emerged when familial sinistrality and strength of left-handedness was controlled for. Thus, when solving the computerized maze test FS+ subjects allocated more time to an initial inspection of the maze test "suggesting

the use of a point-counting strategy assumed to reflect left hemisphere processing" (Levander & Levander, 1990). In the same test strength of left-handednesss interacted with FS with weak FS+ left-handers, especially males, exhibiting particularly low inspection speed and slowness in processing left-sided pathways of the mazes. It was suggested that these subjects "used a problem solving strategy that has been associated with left hemisphere activation and displayed a low level of activation and/or spatial competence of the right hemisphere" (Levander & Levander, 1990). The findings are compatible with the superior verbal performance among FS+ males (Schalling & Levander, unpublished manuscript).

It can be argued from the results that when investigating cognitive performance between handedness groups it is possibly worthwhile to control for sex and familial sinistrality. If solution processes are to be analyzed into separate components (strategy and laterality) it seems as if rather complex tests are required, and that the strength of left-handedness is to be taken into consideration.

Sex, Performance and Problem Solving Strategies

The common assumption of a female superiority in verbal tests and a male superiority in perceptuomotor and visuospatial tests was confirmed in the study (Levander, Levander, & Schalling, 1989). The superiority in tasks requiring skills in verbal functions may be a result of a higher competence of the left hemisphere in females. It is also possible that the left hemisphere becomes more activated regardless of type of task.

In contrast to the very few significant differences in strategy measures between left-handers and right-handers, the results yielded marked differences between the sexes in the mentioned computerized tests. Thus, males appeared to prefer speed over accuracy, and to be more straightforward in their approach to the tests (Levander, Levander, & Schalling, 1989; cf. Klinteberg, Levander, & Schalling, 1987). Females were more cautious in their solving strategies when using the computerized tests. The pattern of sex differences obtained in this study was fully confirmed in an analysis of 400 Mauritian adolescents (15-16 years) who were, at the time, not familiar with computers (Levander, 1987).

Relations Between Handedness, Birth Stress and Cognitive Performance

Assessment of birth stress was included in the study with the intention of elucidating its possible contribution to the origin of left-handedness and its cognitive implications (Levander, Schalling, & Levander, 1989). The assumption was that those left-handers, within a normal population, who had been subjected to birth stress resulting in mild clinical brain damage, would be somewhat inferior in cognitive performance. This would hence be an indirect

link between left-handedness and birth stress. The higher frequency of subjects using the inverted hand position in writing among left-handers would provide further indirect evidence of an association between left-handedness and brain pathology since an inverted writing hand posture has been positively correlated with birth stress among left-handed males (Searleman, Porac, & Coren, 1982).

According to Orton (cited in Hécaen, 1984) problems in language aquisition might be due to an incomplete cerebral lateralization possibly due to birth stress. Consequently a higher frequency of early learning difficulties might be expected, particularly among those left-handers with known birth stress. This would be most prominent in males since they are more susceptible to birth stress than are females.

The results from this large study showed that presence of birth stress was not more common among left-than right-handers regardless of sex, and it was not positively related to inversion in writing hand posture among left-handers (Levander & Schalling, 1988). Furthermore, the findings did show that presence of birth stress was associated with lower cognitive performance, different for males and females, but independent of handedness. However, it was only among females that birth stress was positively related to early learning difficulties and current poor spelling. Among males regardless of handedness, subjects with birth stress were inferior in visuospatial ability as opposed to those without (Levander, Schalling, & Levander, 1989). It thus follows that birth stress and left-handedness may operate independently on cognitive factors and in different ways among males and females.

It has been suggested that left-handedness, immune disorders and cognitive developmental disorders may occur in parallel as a consequence of some common underlying factor. Left-handedness (strong) has been shown to be associated with presence of both immune disorders and learning disabilities (Geschwind & Behan, 1984). Contrary to findings by van Strien and coworkers (van Strien, Bouma, & Bakker, 1987) and Geschwind and Behan (1984) there was no increased incidence of reported early learning difficulties among the 102 left-handers irrespective of sex, and autoimmune diseases were not found to be any more common among left-handers than among right-handers (Levander, Schalling, & Levander, 1989; van Strien et al., 1987).

When subdividing large groups of subjects as was the case in the study referred to above, with the intent of getting more homogeneous subgroups in order to discern consistent patterns in the data more easily, the number of statistical analyses increases, thereby increasing the risk of obtaining significant findings purely by chance. However, in drawing conclusions from the data there are some guidelines for differentiating genuine from spurious significant results. The most important guideline is to interpret the findings in the light of underlying theories which have guided the planning of the study. In addition, the meaningfulness of the results obtained should be considered in terms of their internal consistency and possible replication.

The present sample was large, comprising the major proportion of the population of left-handers attending junior college in the suburb. The size of the sample is important, since left-handers are heterogeneous, which makes it hazardous to draw conclusions when dealing with self-selected groups. In addition, the sample was unselected and heterogeneous with regard to courses choice, since both theoretical and practical courses are given in Swedish junior colleges. This "educational variability" is of theoretical significance, as it may reflect variations in cerebral organization, underlying different cognitive abilities.

However, the relative absence of differences in cognitive functions between left-and right-handers might possibly be a consequence of the sample being a "high ability group" (each sex/handednessgroup receiving a stanine score of 6 or more in vocabulary), which may differ from lower ability samples in patterns of cerebral organization. Consequently, a similar study on a "low ability group" would of course be of interest for further studies.

CONCLUDING REMARKS

Intelligence, here conceptualized as cognitive functioning level, is an important and challenging aspect of human functioning. One aspect of cognitive functioning that appears to be of particular interest is that of problem solving strategy. Extensive research has for example, demonstrated that specific solving strategies are associated with "vulnerability factors" in normal subjects (af Klinteberg, Levander, Oreland, Aasberg, & Schalling, 1987), and in the development of specific maladjustment behaviors (Robins & Rutter, 1990), or syndromes (for a review see Oreland, 1993). Interestingly, the concept of intelligence has been found to act as a moderator in normals on such vulnerability factors (von Knorring, Oreland, & Winblad, 1984).

REFERENCES

Allen, M. (1983). Models of hemispheric specialization. *Psychological Bulletin, 93*, 73-104.

Annett, M. (1985). *Left, right, hand and brain: The right shift theory*. Hillsdale, NJ: Lawrence Erlbaum.

Benton, A.L. (1987). Thoughts on the application of neuropsychological tests. In H.S. Levin., J. Grafman, & H.M. Eisenberg (Eds.), *Neurobehavioral recovery from head injury* (p. 112). New York: Oxford University Press.

Brooks, N.D. (1987). Measuring neuropsychological and functional recovery. In H.S. Levin., J. Grafman, & H.M. Eisenberg (Eds.), *Neurobehavioral recovery from head injury* (pp. 57-60). New York: Oxford University Press.

Cohen, G. (1983). *The psychology of cognition*. New York: Academic Press.

Cook, T.D., & Campell, D.T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand Mc Nally.

Cope, D.N. (1990). Pharmacology for behavioral deficits: Disorders of cognition and affect. In R.L. Wood (Ed.), *Neurobehavioral sequelae of traumatic brain injury* (p. 262). London: Taylor & Francis.

von Cramon, D.Y., & Matthes-von Cramon, G. (1990). Frontal lobe dysfunctions in patients-therapeutical approaches. In R.LI. Wood & I. Fussey (Eds.), *Cognitive rehabilitation in perspective* (p. 165). London: Taylor & Francis.

Dikmen, S., & Temkin, N. (1987). Determination of the effects of head injury and recovery in behavioral research. In H.S. Levin, J. Grafman, & H. M. Eisenberg (Eds.), *Neurobehavioral recovery from head injury* (p. 77). New York: Oxford University Press.

Ellis, A.W. (1993). *Reading, writing and dyslexia. A cognitive analysis.* Hillsdale, NJ: Lawrence Erlbaum.

Ellis, A.W., Flude, B., & Young, A. (1987). "Neglect dyslexia" and the early visual processing of letters in words and nonwords. *Cognitive Neuropsychology, 4*, 439-464.

Ellis, A.W., & Young, A.W. (1988). *Human cognitive neuropsychology.* London: Lawrence Erlbaum.

Eskenazi, B., Cain, W.S., Novelly, R.A., & Friend, K.B. (1983). Olfactory functioning in temporal lobectomy patients. *Neuropsychologia, 21*, 365-374.

Eysenck, M.W., & Keane, M.T. (1990). *Cognitive psychology.* Hillsdale, NJ: Lawrence Erlbaum.

Gardner, H. (1983). *Frames of mind-the theory of multiple intelligences.* New York: Basic Books.

Geschwind, N., & Behan, P. (1984). Laterality, hormones and immunity. In N. Geschwind & A.M. Galaburda (Eds.), *Cerebral dominance: The biological foundations* (pp. 211-224). Cambridge, MA: Harvard University Press.

Gordon, W.A. (1990). *Neurobehavioral sequelae of traumatic brain injury* (p. 176). New York: Taylor & Francis.

Harshman, R.A., Hampson, E., & Berenbaum, S.A. (1983). Individual differences in cognitive abilities and brain organization, part I: Sex and handedness differences in ability. *Canadian Journal of Psychology, 37*, 144-192.

Hécaen, H. (1984). *Les gauchers.* Paris: Presses Universitaire de France.

Herson, M., & Barlow, D. (1976). *Single case experimental designs: strategies for studying behavior change.* New York: Pergamon Press.

Hjaltason, H., Tegnér, R., Tham, K., & Levander, M. (1994). Sustained attention, awareness of disability and ADL assessment in long-term follow up of spatial neglect. Unpublished manuscript.

Kaplan, E., Fein, D., Morris, R., & Delis, D.C. (1991). *WAIS-R as a neurospychological instrument.* Manual. Harcourt Brace Jovanovich.

Kazdin, A.E. (1980). *Research design in clinical psychology.* New York: Harper & Row.

Kinsbourne, M. (1970). The cerebral basis of lateral asymmetries in attention. *Acta Psychologica, 33*, 193-201.

af Klinteberg, B., Levander, S.E., Oreland, L., Aasberg, M., & Schalling, D. (1987). Neuropsychological correlates of platelet monoamine oxidase (MAO) activity in female and male subjects. *Biological Psychology, 24*, 237-252.

af Klinteberg, B., Levander, S.E., & Schalling, D. (1987). Cognitive sex differences: Speed and problem solving strategies on computerized neuropsychological tasks. *Perceptual and Motor Skills, 65*, 683-697.

von Knorring, L., Oreland, L., & Winblad, B. (1984). Personality traits related to monoamine oxidase activity in platelets. *Psychiatry Research, 12*, 11-26.

Levander, M. (1988). Neuropsychology of handedness. Doctoral Dissertation. Department of Psychology, Stockholm University.

Levander, M. (In press). Are there any effects on interhemispheric transfer of information after moderately severe head injury?

Levander, M. (In press). Odor recognition memory in patients with epilepsy.

Levander, M., & Levander, S.E. (1990). Cognitive performance among left-handers differing in strength of handedness and familial sinistrality. *Intelligence, 14*, 97-108.

Levander, M., Levander, S.E., & Schalling, D. (1989). Hand preference and sex as determinants of neuropsychological skill, solving strategy and side preference. *Intelligence, 13*, 93-111.

Levander, M., & Schalling, D. (1988). Hand preference in a population of Swedish college students. *Cortex, 24*, 149-156.

Levander, M., & Schalling, D. (1988). Self-assessed and examinerassessed writing hand posture in swedish left-handers. *Neuropsychologia, 26*, 777-781.

Levander, M., & Schalling, D. (1988). Familial sinistrality—a marker for cerebral bilateralization of language? Superior verbal skill in males with left-handers in the family. Unpublished manuscript.

Levander, M., Schalling, D., & Levander, S.E. (1989). Birth stress, handedness and cognitive performance. *Cortex, 25*, 673-681.

Levander, S.E. (1987). *An automated psychological test system.* Manual. Department of Psychiatry and Behavioral Medicine, Trondheim, Norway.

Levander, M., Tegnér, R., & Caneman, G. (1993). Tactile line bisection in normal subjects. *Perceptual and Motor Skills, 76*, 831-836.

Levin, H.S., Mattson, A.J., Levander, M., Lindquist, C.H.E., Simard, J.M., Guinto, F.C., Lilly, M.A., & Eisenberg, H.M. (1993). Effects of transcallosal surgery on interhemispheric transfer of information. *Surgical Neurology, 40*, 65-74.

Mateer, C.A., & Ruff, R.M. (1990). Effectiveness of behavior management procedures in the rehabilitation of head-injured patients. In R.LI. Wood (Ed.), *Neurobehavioral sequelae of traumatic brain injury* (p. 279). London: Taylor & Francis.

Miller, E. (1992). Some basic principles of neuropsychological assessment. In J.R. Crawford, D.M. Parker, & W. W. McKinlay (Eds.), *A handbook of neuropsychological assessment* (p. 15). Hillsdale, NJ: Lawrence Erlbaum.

Newcombe, F.D. (1987). Psychometric and behavioral evidence: Scope, limitations, and ecological validity. In H.S. Levin, J. Grafman, & H.M. Eisenberg (Eds.), *Neurobehavioral recovery from head injury* (p. 131). New York: Oxford University Press.

Oreland, L. (1993). Monoamine oxidase in neuropsychiatric disorders. In H. Yasuhara, S.H. Parvez, K. Ogushi, & M. Sandler (Eds.), *Monoamine oxidase: Basic and clinical aspects* (pp. 219-247). Utrecht: VSP Press.

Richardson, J.T.E., & Zucco, G.M. (1989). Cognition and olfaction: A review. *Psychological Bulletin, 105*, 352-360.

Robins, L., & Rutter, M. (1990). *Straight and devious pathways from childhood to adulthood.* Cambridge: Cambridge University Press.

Satz, P. (1972). Pathological left-handedness: An explanatory model. *Cortex, 8*, 121-135.

Searleman, A., Porac, C., & Coren, S. (1982). The relationship between birth stress and writing hand posture. *Brain and Cognition, 1*, 158-164.

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

Strauss, M.E., & Allred, L.J. (1987). Measurement of differential cognitive deficits after head injury. In H.S. Levin., J. Grafman, & H.M. Eisenberg (Eds.), *Neurobehavioral recovery from head injury* (p. 88). New York: Oxford University Press.

van Strien, J.W., Bouma, A., & Bakker, D.J. (1987). Birth stress, autoimmune diseases, and handedness. *Journal of Clinical and Experimental Neuropsychology, 6*, 775-780.

Swanson, Y.M., Kinsbourne, M., & Horn, J.M. (1980). Cognitive deficit and left handedness. A cautionary note. In J. Herron (Ed.), *Neuropsychology of left-handedness* (pp. 281-291). New York: Academic Press.

Tegnér, R., & Levander, M. (1993). Word length coding in neglect dyslexia. *Neuropsychologia, 31*,(11), 1217-1223.

CONTEXTUAL AND PERSONAL SOURCES OF INDIVIDUAL DIFFERENCES IN INTELLIGENCE:

EMPIRICAL RESULTS

Rocío Fernández-Ballesteros, Manuel Juan-Espinosa, Roberto Colom and María Dolores Calero

INTRODUCTION

Individual differences in intelligence are among the oldest and most powerful paradigms in the study of the theory, assessment and application of intelligence. In the past few decades, experimental psychology has developed several approaches to the study of intelligence, such as biological correlates (Eysenck & Barret, 1985), cognitive correlates (Hunt, 1978), cognitive components (Sternberg, 1980), cognitive training (Campione & Brown, 1978), and basic intelligence behavioral repertoires (Staats, 1971, 1988). Besides these explicit theories of intelligence, other modern research programs have emphasized the importance of implicit theories (Sternberg, 1990) such as the shared

Advances in Cognition and Educational Practice, Volume 4, pages 221-274.

ISBN: 0-7623-0105-8

knowledge mechanisms used to reduce contextual complexity (Juan-Espinosa, 1994). All of these approaches represent a puzzle and the *Zeitgeist* of the study of human intelligence. Each of them attempts to be the right approach and each of them (apparently) contradicts and criticizes the others.

The issues involved in these approaches differ from one another. For example, the objective from a psychometric scope is: What is the structure of mental aptitudes? The biological correlates approach attempts to respond to: what are the biological bases of intelligence? Cognitive correlates pretend to answer the question: what are the basic processes underlying mental aptitudes? Cognitive components try to respond to the question: which is the right sequence of information processes operating in intelligence tasks? Cognitive training investigates how to improve intelligence performance and predict learning. The behavioral approach explores the environmental conditions which underlie the development of basic behavioral repertoires of intelligence.

In the applied field, the psychometric approach and its mental abilities profile is useful to predict specific job and school performance. Cognitive and component approaches are also useful to predict specific activities required in certain jobs and school tasks. Cognitive training and the behavioral approach are also useful to predict modifiability of intelligence and how it influences school and everyday optimum performance. Biological correlates and implicit theory approaches scarcely have applications at present.

Although different approaches use different experimental or correlational methods (in attempting to reach different objectives), they have in common the utilization of psychometric measurement devices for research purposes (except in the implicit theories approach). Psychometric research uses standard tasks such as letter series, progressive matrixes, and so on. Cognitive correlate investigators predict performance on psychometric tests through cognitive tasks. Cognitive component investigators decompose classic standard psychometric items into process sequences manipulating input presentation. Cognitive training and behavioral research decomposes items from psychometric tests and manipulates instructions, and teaches new cognitive processes in order to train subjects in the basic operations of solving such tasks. Biological correlates search for the associations between biological indexes and psychometric tests. The implicit theorists usually contrast verbal reports with psychometric results.

Is it possible to view individual differences of intelligence from an integrative perspective? In the past fifteen years, several studies on this subject have been conducted in three laboratories[1] of the University Autonoma of Madrid. This chapter outlines the general framework used and the most relevant results obtained.

AN INTEGRATIVE FRAMEWORK OF INDIVIDUAL DIFFERENCES IN INTELLIGENCE

To design an integrated framework of individual differences in intelligence means to deal with human complexity. To take account of this issue in such a framework requires different levels of analysis (Colom, 1995; Juan-Espinosa & Colom, 1989a), ranging from general adjustment person/context to the specific individual/task performance.

As can be seen in Figure 1, the adjustment *Person/Context* can be analyzed as: (a) the *resources* a person can draw on to solve the relevant problems demanded in a specific society, and (b) the *context requirements*, in terms of problems and task *demands*, and the ways (*mediators*) of learning how to develop appropriate intelligence skills for solving problems and performing culturally relevant tasks.

At the general level, the cross-cultural scientific literature on intelligence highlights the idea that different *contexts* set different problems for *persons*. In terms of the *Law of Cultural Differentiation* (Ferguson, 1956), this means that "cultural factors prescribe what shall be learned and at what age; consequently, different environments lead to the development of different patterns of ability" (p. 121). That is, the general "person/context" adjustment can be taken as a trade-off between the personal resources of the individuals and the contextual requirements that a specific society requires for adaptation purposes. Analysis in this field includes the so called "Implicit or Informal Theories" (Sternberg, 1990). In this approach, the goal is to discover the form and content of people's informal theories when they live in different contexts, and to contrast these in order to uncover shared and non shared characteristics in the division of abilities. Some of our studies have been devoted to this (Juan-Espinosa, 1994a, 1994b) and they will be explained later.

From the *Person* starting point, personal resources can be divided into: (a) biological resources, (b) processes resources, and (c) abilities resources. All are part of the whole landscape of the person's intelligence.

Biological resources are the structural powers and limits a person has at a specific moment of his/her life-span. These resources are a necessary condition for the upper levels (i.e., processes and abilities) of personal resources, but they are not sufficient in themselves. It is relatively easy to discover both biological/processes and biological/abilities correlations, but relatively difficult to determine the causal relationship between a given biological measure and a given intelligence score. In any case, it is clear that biological resources should be taken as *biological correlates* of intelligence. The main kinds of biological correlates are (Juan-Espinosa, in press a; Juan-Espinosa & Colom, 1989b): *physical, anatomical, functional, and molar*. The level of auditive and/or visual impairments are examples of sources of individual differences in physical correlates; brain size, as an example of the anatomical correlates; and finally,

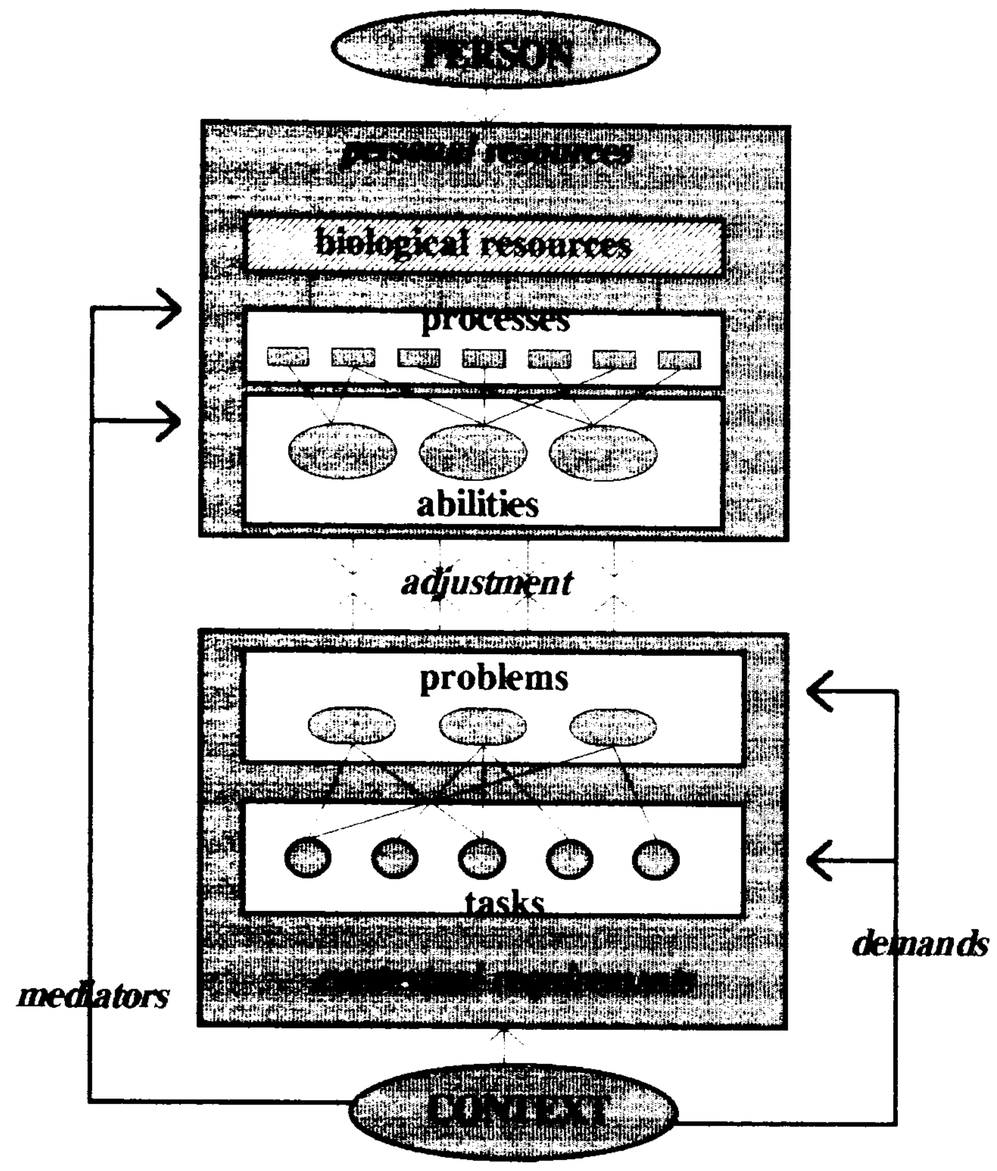


Figure 1. Framework of Individual Differences Sources in Intelligence and Learning

velocity of nerve conductance, evoked potential responses, and so forth, are examples of the functional correlates. In biological resources research, biology is also manipulated as a molar variable by means of medical exams; subject classification as "organic/non-organic" is an example of this type of source of individual difference in the molar biological correlates. Some of our work has been devoted to the study of physical sources (Blanco et al., 1992), functional sources (Icarán et al., 1993), and molar sources of individual differences (Fernandez-Ballesteros & Calero, 1992, 1993).

Cognitive *processes* resources reflect how the person behaves while performing the specific tasks involved in the problems of a given context. These resources can be divided into *cognitive correlates* and *cognitive components*. While the first try to establish which basic processes underlie cognitive abilities, cognitive components try to establish how an individual behaves cognitively when performing a task; that is, they attempt to establish the sequence in which the processes operate. Several authors have explored cognitive resources as sources of individual differences. Kyllonen and Christal (1989) proposed *mediator* and *enabler* classifications of four sources of individual differences (i.e., capacity of working memory; execution speed; the breadth of declarative knowledge, and the breadth of procedural knowledge). Campione et al. (1982) introduced *accessibility* and *availability* of declarative and procedural knowledge as sources of individual differences. Sternberg (1980) added *componential sequences* as an

individual difference factor in cognitive performance. Juan-Espinosa (Juan-Espinosa & Colom, 1989a; Juan-Espinosa, In press a, In press b) added the *modulator* source, taking into account the differences produced by the levels of spontaneous adaptation of a *mode* (Verbal or Visual-Spatial) of representing information in the working memory. This has been one of our main focuses of research in the Laboratory of Individual Differences at the Universidad Autónoma of Madrid. The modulator source requires further comment.

One way of thinking about this issue is by looking at the *working environment* computer systems. The working environment refers to the material we work on to process the information. *Graphics* and *Text* are the two main computer working environments or *modes* for dealing with information. Their properties differ: *Graphics mode* works with dots or pixels (307.200 approximately), while *Text mode* works with positions of characters (2000 spaces approximately). Both can work with images and texts, but graphics is better suited to dealing with figures and colors of arbitrary complexity, while text is better suited to working with letters and numbers. In the computer metaphor, these modes are related to the main, broad abilities Crystalized Intelligence and Visual-Spatial ability, found in the main psychometric models. Our main concern at this point is the relationship between modes and abilities, and how different modes affect attention, memory, and performance on reasoning tasks (Juan-Espinosa, in press b).

Finally, cognitive *abilities* resources reflect the individual tendencies of consistent general performances on types of *problems*. The main sources in these resources are the personal profiles of abilities and the level reached in each ability at a specific moment in the life-span. Levels of General Intelligence (*G*), Crystalized ability (*Gc*), Visual-Spatial ability (*Gv*), and so on, are some examples of these sources of individual differences. Various studies of the relationships of this kind of resource to biological and cognitive resources, either in performing tasks or in learning intelligent skills, have been carried out by the authors of the present chapter.

On the other hand, taking *Context requirements* as a reference, problem and task demands and social mediators of learning can be taken as the main sources of individual differences at the *Context* level, whether micro or macro.

As demand requirements in a micro-context level, *problem* demands point to the types of problems involved in a specific context. Some of them require the same type of problems for example, journals require working with words expressed verbally or in written form; publicity requires working with shapes and forms; others require two or more different types. *Task* demands reflect the different tasks needed to accomplish types of problems. Letter classification by Zip Code in a post office is an example of a task.

Also, the macro-context can be derived from the social perspective as a mediator in learning processes. That is, context refers not only to which problems and tasks are relevant for adaptation purposes (demand

requirements), but also how to solve them. Based on this approach, different societies provide different kinds of institutions (e.g., schools, universities, etc., in western culture, or House of Word in some African cultures) and people (e.g., fathers, teachers, etc.) who are in charge of transmitting cultural values, and teaching relevant skills. Studies on the role of social mediator in skills training have been carried out by the Fernández-Ballesteros research team.

PERSON/CONTEXT ADJUSTMENT

The study of intelligence in different contexts may be approached either from the point of view of how people solve problems, or from the opinions they voice about intelligence as a concept in itself and/or as a property of people living in that environment. In the first case, scientists develop and expound theories through a network of hypotheses about intelligence, whether they be general or differential, which they try to test using appropriate methods and techniques. This approach is called "Explicit Theories" or "Formal Theories." Alternatively, when attempts are made to uncover the conceptual network underlying opinions about intelligence, one is said to be using the "Implicit theories" approach. In the formal theories approach, concepts relating to intellectual functioning are useful scientific "inventions" or scientific constructs (Eysenck, 1979); implicit theories are people's constructions (whether by scientists or lay people) that reside in the mind of these individuals, which scientists try to "discover" (Sternberg et al., 1981).

Research on Implicit Theories of Intelligence

When we study intelligence from the implicit theories approach, the data we are interested in are people's communications (in whatever form) regarding their notions of the nature of intelligence as tied to their ecocultural contexts. In this sense, people's concepts of intelligence have the aim of reducing the complexity and variability of their own and others' problem-solving behavior to a structure of limited concepts that allows them to categorize people by prototype. We (Juan-Espinosa, 1994a) have, therefore, considered the implicit theories that involve the structural and functional interlacing which underlies the mental ability concepts and which makes them useful as mechanisms of knowledge about the world and about ourselves. Our idea, in this sense, is that implicit theories of intelligence, as a mechanism of knowledge, may be of a quasi-universal nature, and that their attributes are organized in a diffuse hierarchy in which the globality-molecularity continuum of their components or properties reflect the degree to which groups and/or individuals are immersed in particular cultures, according to the above-mentioned law of cultural differentiation.

Berry (1987) proposed some theoretical mechanisms for the law of cultural differentiation. One of them, acculturation, focuses on the issue of how problem-solving behavior is affected by cultural influences whithin the sociopolitical context of one's own group. It may be reflected by the presence of schooling and urbanization, among other forms. Our specific interest in this field was to uncover how acculturation influences the structure and content of people's conception of intelligence.

To accomplish this aim, we focuses in our study on an ex-colony of Spain: Equatorial Guinea, an African country located between Cameroon and Gabon. There are 387,000 inhabitants occupying 28,051 square kilometers which are distributed between the continental and island areas, Bioco being the main island and the location of Malabo, the political capital. The main ethnic groups in the continental area, where the research took place, are *Combes, Bujebas* and *Fang*. Our sample was entirely composed of Fang people. The Fang are a sub-ethnic Bantu group and the majority ethnic group in the continental area of Equatorial Guinea. Until only very recently, they had no written language. Even today, the use of writing is not common among the population. The official language is Spanish, though it is hardly spoken among the rural population. The Fang have a clan-based social structure. Marriages are exogamic and patrolinear.

In order to study the acculturation influence, two ecocultural contexts were selected: (a) A traditional/rural one, represented by *Nsoc* (2,345 km^2), an isolated north-eastern region near the Cameroon and Gabon borders, and (b) an acculturated/urban one, represented by Bata City, the financial capital of Equatorial Guinea.

The entire research project was carried out using the Fang language. It was undertaken in two stages plus a preliminary stage. In the preliminary stage, we determined whether quantitative terms were used to compare people, and if people's intelligence was assessed in any systematic way. The next stage was (1) to discover the general properties of people's implicit theories of intelligence in each context, and (2) to obtain a master list of characteristics of an intelligent person. The last stage was a quantitative valuation of these characteristics uncovering the structure and content of people's implicit theories in both contexts.

The *preliminary stage* was undertaken in Bata City. In order to carry out the objectives, we collaborated with a bilingual Fang/Spanish anthropologist, a native of Equatorial Guinea.

The Fang language has one word to describe an intelligent person: *nfefeg*. It is interesting to note that *nfefeg*, or "intelligent," is derived from the word *feg*, which can be translated as: intelligence, rule, or vision. The last item is particularly interesting when compared with the opposite of intelligence, *ndjimam*, which comes from the word *ndjibe*, meaning darkness. In the Fang culture, an intelligent person or *nfefeg* is someone with *achi* or profound

thinking, a person who can see beyond the superficial appearance of things, while a *ndjimam* is a person who only sees outside appearances ("s/he is in the dark").

In addition to the word *nfefeg*, which is used to describe an intelligent person, there is also the word *mbomam*, used to describe a smart and skilled person who gets on well in everyday life, and the word *eyemam*, used to describe a knowledgeable person or someone capable of knowing and learning things. The opposite term to *eyeman* is *eyeyemam*. *Mbomam* does not have an opposite term.

In the first stage, the objectives were: (1) to uncover the general properties of the implicit theories concerning the intelligent, knowledgeable, and skilled person in the two different contexts, and (2) to obtain a master list of behaviors that were characteristics of the intelligent person. This stage involved 218 subjects, including 94 (50 males and 44 females) living in villages in the *Nsoc* area (traditional context) and 124 (72 males and 52 females) living in Bata City (acculturated context). The mean age was 32, ranging from 19 to 54 years old.

Judging by the frequencies of behaviors listed for each type of intelligent person in each of the contexts, we may conclude from this phase that people in the traditional/rural context perceived the intelligent person as being substantially similar to the knowledgeable ($r = .4851$ $p < .001$) and skilled ones ($r = .4461$ $p < .001$); and that being skilled and being knowledgeable were also related, but at a less significant level ($r = .3011$ $p < .05$). In contrast, people in the urban/acculturated context perceived the intelligent person not only as substantially more similar to the knowledgeable one than people in the traditional context ($r = .7486$ $p < .001$), but saw these as the only related concepts, since there were no other correlations with or among the remaining concepts under research.

Taking correlation as an indicator of shared concepts between people living in each context, we conclude that the most shared concept was that of the intelligent person ($r = .5845$ $p < .001$), the skilled ($r = .3817$ $p < .001$) and knowledgeable ($r = .2956$ $p < .05$) concepts being the least shared ones.

In the *second stage* we sought to uncover the conceptual structure underlying the idea of an intelligent person in each selected context.

One hundred thirty five (81 males and 54 females) and 148 (88 males and 60 females) adults (not the same subjects as in the first stage) from the traditional and acculturated contexts were interviewed. The mean age was 34 and ranged from 17 to 56 years.

In order to analyze the structure and underlying content of the concept of the Intelligent Person in the traditional/rural and acculturated/urban contexts, two factorial analyses were performed on the checklist of 46 characteristic behaviors. Main Components method and Varimax rotation were used in both analyses. Behaviors with a loading of .50 and above were taken into account for interpretation purposes.

Table 1. Main Characteristic Behaviors Loading Each Factor (Traditional/Rural Context)

N° FACTOR I: Social competence	
1.	To orient people properly.
35.	To do things for the good of others.
6.	To give good ideas.
31.	To solve the problems in the village.
21.	To give good advice to others.
8.	To solve conflicts among people.
N° FACTOR II: General reasoning	
2.	To interrelate things among themselves.
28.	To think well and methodically.
34.	To see problems before they appear.
38.	To reason well.
16.	To solve difficult problems.
N° FACTOR III: Visual-Spatial ability	
7.	To imagine how to do things before doing them.
4.	To imagine things before doing them.
27.	To know how to go into the jungle and not get lost.
36.	To know how to build houses and to build them.
N° FACTOR IV: Verbal ability	
13.	To know what you are saying.
45.	To know what a person is going to say before s/he speaks.
15.	To be convincing when talking.
8.	To solve conflicts among people.
N° FACTOR V: Learning ability	
10.	To be patient when thinking.
30.	To learn things properly.
40.	To know how to bear suffering in order to get something.
26.	To know how to organize things.
N° FACTOR VI: Traditional-verbal ability	
11.	To explain events through stories and sayings
43.	To have a good memory.
39.	To speak well.
17.	To speak in such a way as to be understood.

A rotated factorial structure was obtained in the *traditional/rural context*. Six factors explained 55.5 percent of the variance. These factors were labelled *Social competence* (11%), *General reasoning* (10%), *Visual-Spatial ability* (9%), *Verbal ability* (8.6%), *Learning ability* (8.5%), and *Traditional-verbal ability* (7.7%). Table 1 shows the main behaviors loading each of the six factors respectively.

Table 2. Main Characteristic Behaviors Loading Each Factor

N°	
	FACTOR I: Social competence
34.	To see problems before they appear.
15.	To be convincing when talking.
5.	To know how to economize.
37.	To know how to negotiate with money.
21.	To give good advice to others.
8.	To solve conflicts among people.
26.	To know how to organize things.
1.	To orient people properly.
14.	To treat others well.
41.	To know the intentions of others.
42.	To know how to distribute among others.
16.	To solve difficult problems.
38.	To reason well.
9.	To know how to observe things.
10.	To be patient when thinking.
	FACTOR II: Verbal-Educational ability
45.	To know what a person is going to say before s/he speaks.
13.	To know what you are saying.
23.	To know writing.
46.	To know how to play with numbers.
29.	To know reading.
40.	To know how to bear suffering in order to get something.
39.	To speak well.
	FACTOR III: General reasoning
28.	To think well and methodically.
3.	To plan things.
16.	To solve difficult problems.
38.	To reason well.
18.	(Negative) To know how to plant and when to harvest.
11.	(Negative) To explain events through stories and sayings.
	FACTOR IV: Memory and learning ability
43.	To have a good memory.
30.	To learn things properly.
44.	To learn quickly.
9.	To know how to observe things.
	FACTOR V: Visual-Spatial ability
7.	To imagine how to do things before doing them.
27.	To know how to go into the jungle and not get lost.
4.	To imagine things before doing them.
	FACTOR VI: Cautiousness
24.	To be able to stop evil coming into the home.
35.	To do things for the good of others.
10.	To be patient when thinking.

Again, a rotated factorial structure was obtained in the acculturated/urban context. Six factors explain 69 percent of the variance. These factors were labelled *Social competence* (22%), *Verbal-Educational ability* (11.6%), *General reasoning* (11%), *Memory and Learning ability* (8.7), *Visual-Spatial ability* (7.7), and the last one, though difficult to describe, we labelled *Cautiousness* (7.5). Table 2 shows the main behaviors loading each of the six factors respectively.

Judging by the factors obtained and their characteristic behaviors included in each context, we learned:

First, most of the factors obtained for the intelligent prototype in both contexts overlap considerably, as would be expected from the simple correlations made in the first stage between contexts, but they also differ in some ways. The order of appearance of factors differs, probably reflecting their order of importance for differential adaptation purposes. Content in each factor also differs, probably according to how each ability takes shape in relation to the ecological demands of the skills in each context. Finally, the two contexts seem to demand two different specific culture-bound factors: *Traditional-verbal ability* in the rural and *Cautiousness* in the urban context.

Second, as far as the factors themselves are concerned, some of them are outwardly quite similar to the main factors obtained in formal psychometric theories (Carroll, 1993; Cattell, 1971, 1987; Gustafsson, 1988; Vernon, 1969, 1979). *General reasoning, Verbal* and *Verbal-Educational ability*, *Memory* and *Learning ability*, and *Visual-Spatial ability* mirror the "*g*" and *Gf* factor, *Gc* and *V:e*d factor, *2Y* (General Memory and Learning in Carroll's theory), and *Gv* factor respectively. On the other hand, *Social competence*, though proposed as "Social Intelligence" by Thorndike (1920) and missing in explicit psychometric theories (Keating, 1978), can be found in some implicit theory research (Sternberg, 1990), and plays the main role in people's prototype of an intelligent person from both contexts in our research.

Research on Implicit Theories of Spatial Orientation

This study explores people's ideas about how they maintain their orientation in large-scale space. The Fang-okak hunters of Equatorial Guinea, a jungle-dwelling society, were taken as an example of expert ground navigators. Fang-okak hunters were asked to make judgments of the characteristics of people with a high level of ability in spatial orientation, of situations that require this ability to a greater or lesser extent, and of the series of actions or strategies necessary to guide oneself to a known place and then to an unknown place.

The ability to know where we are, where things are in space, and to follow paths is termed "Spatial Orientation Ability." Although all human beings possess this ability to some degree, it appears that there are strong individual differences in the precision with which people can place themselves and other objects or people in the environment.

Spatial Orientation Ability (SO) is one of the classic subjects of study for Explict Theories in Psychology. From the point of view of Individual Differences, SO has been identified in many factor analyses, beginning with the work of Guilford and Lacy (1947). However, its identification as an independent ability has been questioned on a number of occasions, especially with respect to visual-spatial ability (Vz) (Borich & Bauman, 1972). In spite of this and from a generic point of view, spatial orientation ability may be seen as an underlying ability of Visual-Spatial Aptitude (Gv) (Lohman, 1988; Pellegrino & Hunt, 1991). Interesting tests have been developed that allow us to come closer to understanding the strategies people use for working with this ability (Guilford & Zimmerman, 1947; Hoffman et al., 1968; Eliot & Smith, 1983).

Another means of approaching the study of orientation ability is to take the implicit theories about it and the task of orientation as a point of reference. In this sense, we tend to voice opinions about this ability. These opinions are: (a) as much about ourselves as about other people, (b) about the environment and situations, suggesting that in certain places it is more difficult to orient oneself than in others, and (c) about the actions needed to orient oneself according to the circumstances (Juan-Espinosa, 1994b). These three types of judgment may reflect an implicit theory about spatial orientation developed by the individual.

In this study we (Juan-Espinosa, 1994b) attempt to approach the problem of orientation ability by taking, as a starting point, people's judgments about orientation. The aim is to find out what the underlying characteristics are when we make a judgment about this ability, whether the judgment be about another person, the terrain, or the actions or strategies necessary for orienting oneself.

The Fang people inhabit the interior zones of the equatorial rain forests (the ecosystem with the greatest biomass in the world). A homogeneous annual temperature (mean 33° Centigrade) and high levels of annual precipitation 2,000 mm) produce lush vegetation. It is very dark inside the rain forest, the almost 100 percent coverage by vegetation and the several levels of tree vegetation (their crowns forming roofs more than 30 metres high) constitute a real obstacle for external visual references. The vegetation of the rain forests is so dense that *not only the sky references, but also the references from internal landmarks are greatly reduced.*

All of the research was carried out using the Fang language. It was undertaken in two stages, plus a preliminary stage. In the preliminary stage, we determined whether quantitative terms were used for comparing people, and if people's orienting ability was assessed in any systematic way. The next stage was to: (1) discover the general properties of the orientation task with respect to people, situations and actions enabling one to guide oneself to a known or an unknown place, and (2) to obtain a list of characteristics of (a) the person, (b) the situation and (c) the strategic actions taken in orientation. The third phase was to ask for a quantitative valuation of these different characteristics.

In the preliminary stage, the result obtained was that the Fang language has no one word to describe spatial orientation. However there is a character who appears in fables and tales who never finds himself lost wherever he goes, and this "gift" is considered socially desirable. The nearest word for describing a person with these characterics is *nfefeg* (this word has other meanings too, as we have seen above), or intelligent. The person who never gets lost, or is always able to find the "right path" is *mbot abele feg akoan mben a zen* or "a person with the intelligence to find the right path."

The aims of the *first stage* were: (a) to reveal the general properties of the implicit theories concerning the orientation task, and (b) to obtain a list of behaviors, situations and strategies or actions that are characteristic of orientation in the natural wider environment, thus permitting us to enter into the second stage.

28 males were interviewed, seventeen of whom were expert adult hunters and eleven adolescent apprentices. They were all "fang okak" and inhabitants of the nearby villages in the Niefang Mountains.

The results obtained were that, while the two groups, experts and apprentices, share a concept of orientation ability ($r = .4963$ $p < .05$), this concept is not identical for the two groups. This is probably due to the fact that experts are exposed to a wider range of situations in their work where this ability is required. It might also be the case that selection has "creamed off" those individuals with greater ability. This appears more evident in the self-ratings: while experts tend to think that orientation ability is more related to intelligence ($r = .5927$ $p < .05$), apprentices tend to associate it with smartness in everyday life.

Focusing now on situations, there is, as expected, a significant negative correlation between situations of high and low difficulty ($r = -.7886$ $p < .05$). The relationship between medium difficulty and high and low difficulty is not significant, though it is negative in relation to high difficulty and positive in relation to low difficulty.

The relationship between actions to guide oneself to a known place and to an unknown place appear to be different, in general. Moreover, it appears that experts and apprentices share a theory about ideal courses of action to get to a known place ($r = .9036$ $p < .001$), but not to get to an unknown place ($r = -.0375$). This may be because both groups have a lot of experience of getting to known places, due to the custom of visiting nearby villages, while it is the hunters who are principally exposed to orientation tasks in unknown places, because of changes in animal habitats and continual changes in the rain forest landscape.

The aim in the second phase of research was to discover the structure underlying the idea (a) of a person who is intelligent in finding the right path, (b) of situations judged as being related to orientation, and (c) of behaviors or strategies found in people with a high and low orientation ability in the task of getting to a known and an unknown place.

Twenty six adults (others than the subjects of the first stage) were inteviewed. All of them were male hunters, belonging to the Fang-okak clan and also inhabitants of villages situated on the slopes of the Niefang Mountains.

In order to analyze the structure and underlying content of the concept of orientation ability, 24 characteristic behaviors obtained in the first stage were factor analyzed. The Main Components method was employed in this analysis, in order to obtain the factorial matrix without rotation and the oblique rotation "Oblimin Direct" with gamma 0. The result was that 7 factors were given with a proper value that was higher than unit value. Taken as a whole, the seven factors explained 71.85 percent of the total variance.

To analyze the structure and content of the expert hunter concept of (a) a situation related to the orientation task, and (b) actions or strategies applied to get to an unknown and known place, the cluster analysis technique was performed.

At the end of this stage, we learned:

A. That the Fang-okak hunters believe that the features (i.e., the factors obtained) of a person with a high orientation ability are: anticipation through mental images of places and situations; discrimination of meaningful signs in the environment; interpretation of body signals, being able to separate oneself from the surroundings in order to note the sensations of gravity and recognize the unevenness of the land, and to make time estimates and route estimates; self-control of emotional states and episodic memory, which implies the use of experience to solve problems; declarative knowledge of landmarks, knowing how to use the fixed and most prominent elements of the land as points of reference; and declarative knowledge of indicators, being able to obtain clues about the situation.

B. that the situations in which individuals have the opportunity of using these abilities would seem to be, in the first place, those where the field of vision may be limited either by meteorological factors, such as fog, or by biogeographical factors, such as the height and density of the vegetation. This type of situation is judged as being the most difficult for orientation; in the second place, those situations where one has to guide oneself by both human and animal indicators; and in the third place (judged as being of least difficulty) those situations that allow for a view of the terrain, thus allowing one to set out on the best path towards the given destination.

C. And that finally, the task of getting to a known or an unknown place appears to require a range of actions, from basic strategic actions that anyone could carry out, such as anticipating the necessities of the route (known and unknown place) or finding the right direction (unknown places), to that of more complex actions in which mental images play an important role in the planning of places and paths (known place, unknown place) and the recognition of destinations (unknown places), as well as in the basic actions for the development of analogical inferences that allow one to guide oneself to

unknown places.

In relation to formal theories, the most characteristic features of spatial orientation seem to coincide (a) with the psychometric perspective, in the sense that Spatial Orientation is not a unitary ability, but a popular concept that implies a variety of abilities such as Visualization, Spatial Relations, and so on (Carroll, 1993; Lohman, 1988) and (b) with the differential abilities suggested by Hunt (personal communication; Infield et al., 1990) based on the model by Siegel and White (1975) for large spaces and work by Downs (Downs, 1981; Downs & Stea, 1977) 'Observational Ability, Imagery Ability and Inference Ability', but with the peculiarity that the first factor in this present work is that it unites imagery ability and inference ability. However, the factorial results point to another ability which could be important if it were to be confirmed by further studies, the *Ability to notice body states*, which, in part (only in part), seems to coincide with Witkin's (1967) "Field-Dependence-Independence" construct, since it includes characteristics such as noting sensations of gravity. Declarative knowledge about landmarks and declarative knowledge about indicators are also present in other research as a method of updating one's position (Baker, 1981; Etienne, 1992) but, especially in the case of indicators, such signals are not only visual but also audible landmarks (Loomis et al., 1993), something not found in our research. We have not been able to find a reason for this absence, but the small sample used for drawing up the master lists of characteristics is a possible justification.

In relation to other informal theories research, implicit theories of Spatial Orientation seem to be part of the Fang's concept of intelligence and, specifically, part of *Visual-Spatial ability*, as found in the work of Juan-Espinosa, (1994a). Variables such as "To imagine how to do things before doing them," "To imagine things before doing them," "To know how to go into the jungle and not get lost," and "To know how to build houses and to actually build them" were the main characteristic behaviors, with loading of .50 and above in this factor.

PERSONAL RESOURCES AND MICROCONTEXTUAL DEMANDS

The relationship between the person and the contextual requirements can (and must) be studied by designing well-known tasks. In our general frame of reference (see introduction) such tasks pose microcontextual demands (or requirements).

At the UAM Individual Differences Lab., we are exploring personal characteristics by analyzing their global verbal and spatial processing resources. Verbal and spatial abilities are supposed to be two of the more important cognitive tools of human intelligence. So, by hypothesis, these two cognitive

tools *mediate* the performance patterns of subjects who have been asked to solve a variety of tasks requiring intelligent behavior.

This conceptual hypothesis can be derived from the fact that analyzing the components of information processing in order to explain intelligent behavior is a very unordered process (Snow, 1989; Snow & Lohman, 1993; Kyllonen et al., 1984a, 1984b). This means that components of information processing are particular to the person who is solving the various tasks (in other words, components or combinations of components are very idiosyncratic).

However, one proposed solution is, to *analyze cognitive strategies* (Cooper, 1982; Cooper & Mumaw, 1985; Hunt, 1980; Lohman, 1988). We know that a key issue related to cognitive strategies is the *type of mental representation* selected by different individuals to code the information derived from the tasks, in order to operate and to respond to them. Curiously, when we pay attention to the type of representation, it is easy to find a relevant order that increases the organization of the *individual differences in cognitive strategies* (Colom, 1995; Juan-Espinosa, in press a).

This leads us to conclude that different individuals might express their preference for certain kinds of mental representation based on their cognitive resources for processing verbal or spatial information; that is, the major categories of information in intelligence tests (Guttman, 1965; Snow et al., 1984). These resources might mediate the configuration of a particular cognitive strategy, and must aid the organization of individual differences.

On the other hand, one way of exploring cognitive resources is to administer psychological tests which are factorially loaded in the verbal and spatial abilities or in the cognitive tools mentioned above. Let us now look at the experimental framework for contrasting the conceptual hypothesis just expounded.

Experimental Framework

Following Cooper (1982), Sternberg (1980), Lohman (1988), Hunt (1980), Lohman and Ippel (1993) and others, experimental subjects are classified on the basis of their scores on two tests with high loadings in verbal and spatial factors (Factor V from the Primary Mental Abilities, and Identical Figures from the Factorial Abilities Battery). Threshold levels are scores above the 75 percentile for high ability, and scores below the 25 percentile for low ability.

Performance on those tests allows a *sampling* of the verbal and spatial cognitive resources required to solve a variety of tasks of different natures (complexity and content; Guttman, 1965; Snow et al., 1984). Yet the classification criteria are extreme; it is not necessary to use large samples (comparison of extreme groups strategy, see Colom, 1995; Cooper, 1982; Juan-Espinosa & Colom, 1989b).

These are the experimental groups:

1. Subjects with high scores on both tests. In this case, the hypothesis is that they have good verbal and spatial cognitive resources (high abilities group).
2. Subjects with high scores on the verbal test, but low scores on the spatial test. The hypothesis is that they have good verbal but poor spatial cognitive resources (*verbalizers*).
3. Subjects with low scores on the verbal test, but high scores on the spatial test. In this case, the hypothesis is that they have poor verbal but good spatial cognitive resources (*visualizers*).

Performance patterns for tasks of varying complexity and content are analyzed by virtue of the nature of the cognitive resources of the different experimental groups. The variables are usually measured next: response times (hit trials), accuracy levels, and in some cases, examination of the basic material contained in the tasks. Performance patterns and measured variables lead to heavy ideas about the cognitive strategies that the subjects appear to use.

Tasks are manipulated by varying the cognitive demands of the resources in the *spheres* of attention-perception, working memory, and reasoning. The tasks range from those demanding very few cognitive resources to more complex tasks which, therefore, require more cognitive resources and a greater mental effort.

Individual Differences in Attentional and Perceptual Resources

Prior to selecting the relevant tasks or experimental paradigms, the basic questions are: which kind of stimuli are attended with more efficacy? Can the subject be apart from the meaning of a word (or a purely visual variable like color) when they have to perceive speedily and precisely?

Selected experimental paradigms are the classic stroop task and the sentence picture verification task, well known in individual differences research, and both demanding verbal and spatial processing resources.

Stroop Task

In the classic stroop task, subjects must give the color of a given written word (GREEN), or give a word incongruous with the color (RED). Furthermore, stimuli can be projected to the left or to the right hemisphere. This allows the analysis of certain benefits in terms of processing speed. Subjects must hence give either the color or the word, and the stimuli can be tachitoscopically projected to the left or right hemisphere. The available data leads us to conclude that subjects have more trouble naming the color and ignoring the word (MacCleod, 1991).

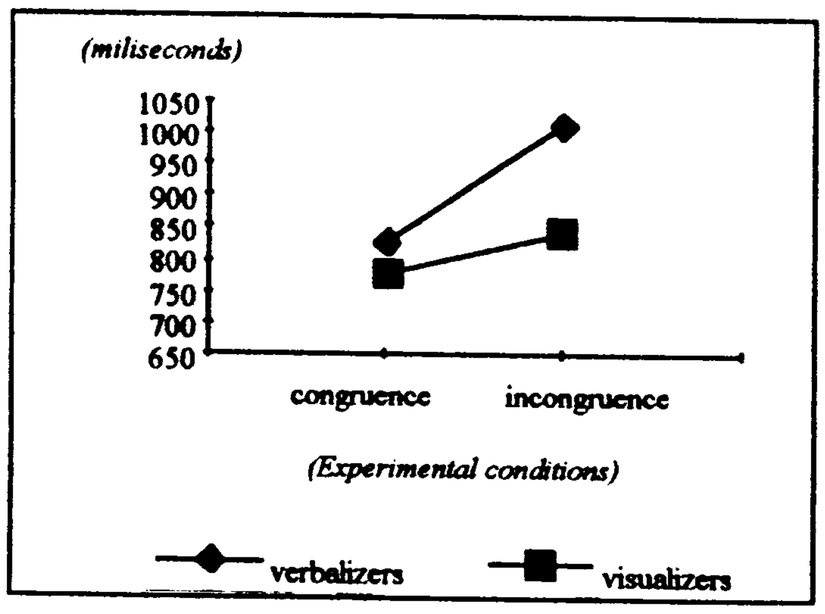


Figure 2a.

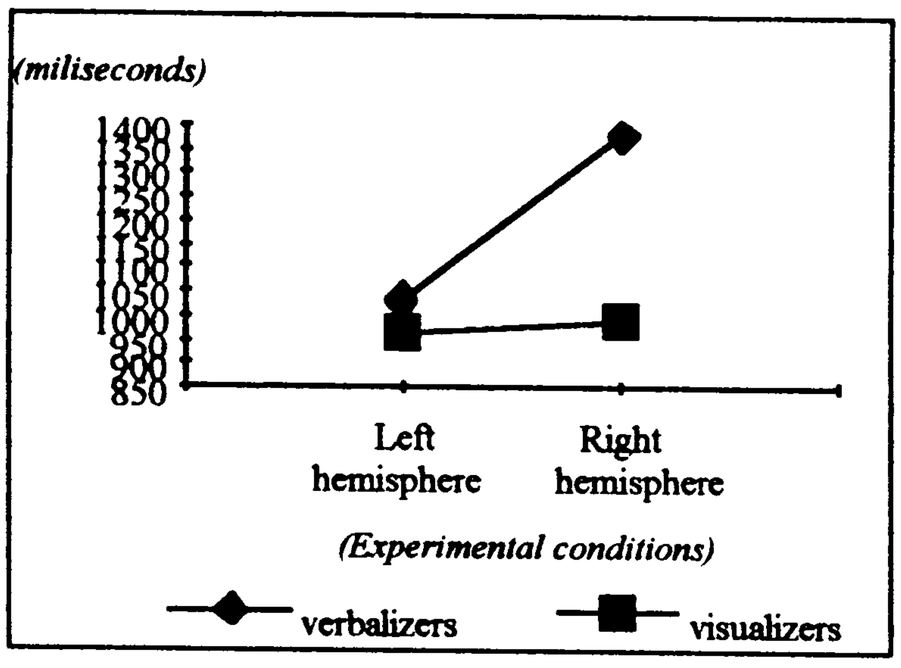


Figure 2b.

We explored two processing levels (congruence and incongruence between color and word) involving two experimental groups (verbalizers and visualizers) and two conditions of stimuli lateralization (left hemisphere and right hemisphere). There were forty-seven stimuli, and each experimental group had five extreme subjects.

The data show that the stroop interference effect is much more pronounced for verbalizers than for visualizers (Blanco, 1994; Blanco & Colom, in press): (a) response time is higher for verbalizers than for visualizers when color and word are incongruent (see Figure 2a); (b) the same temporal pattern applies in the color-naming condition; and (c) when stimuli were projected to the right hemisphere and subjects asked to name the color, visualizers retained their response time, but the time for verbalizers increased dramatically (Figure 2b).

The verbalizers' processing speed decreases when demanding tasks (naming color) and the hemispheric condition (right hemisphere) were coincident with their poor cognitive resources as sampled by a mental test.

Sentence Picture Verification Paradigm

A simple sentence ("Plus above star") and then a picture representing a figural presentation of the elements belonging to the sentence are shown on a screen:

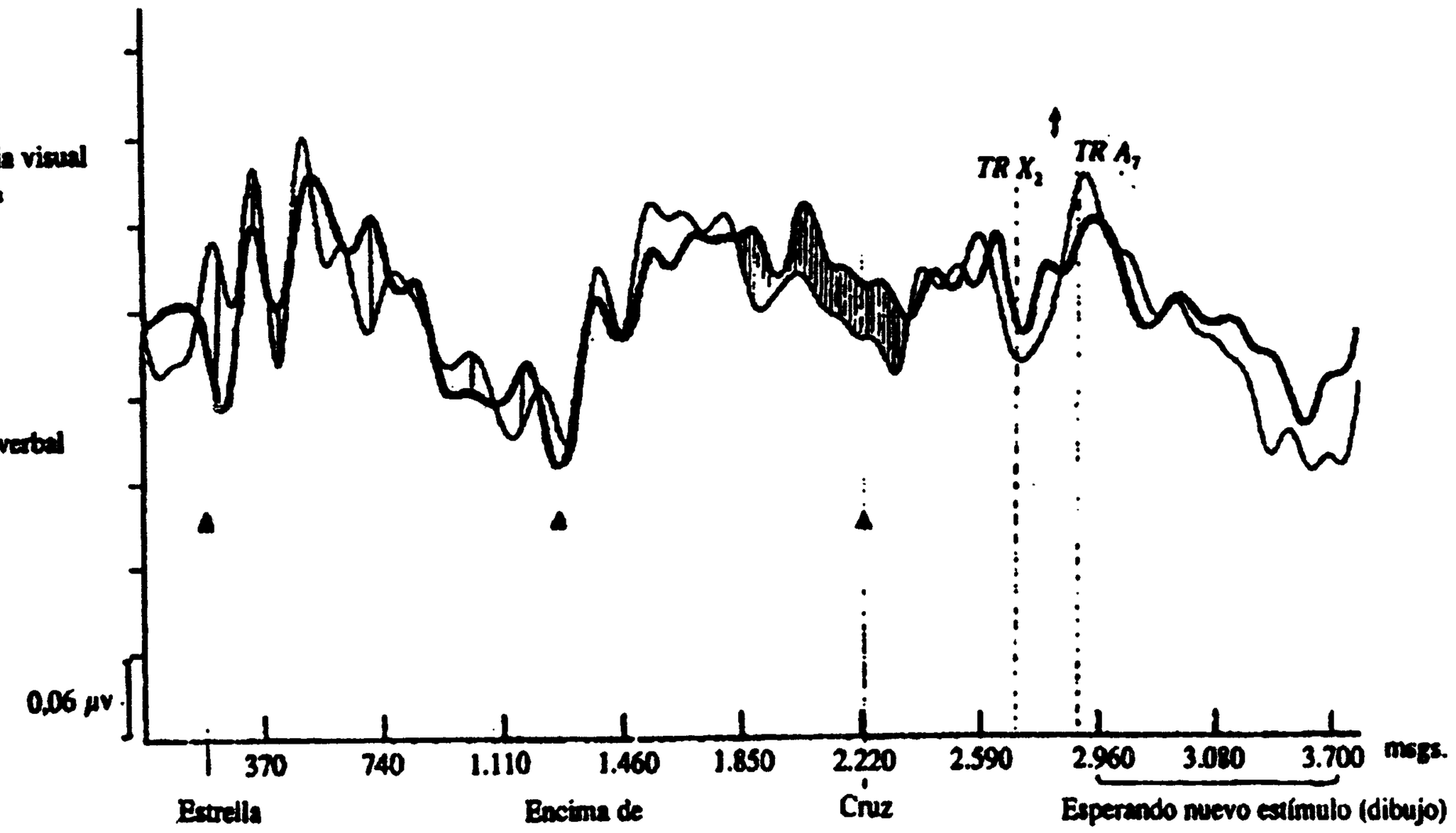


Figure 3.

+
*

Subjects must verify the figural information. Pure linguistic complexity is one of the most common criteria for analyzing this task (Just & Carpenter, 1976): true affirmative, false affirmative, false negative, and true negative.

However, individual differences research has shown that some subjects transform the sentence into a mental diagram, or the picture into a sentence. This leads to *different processing models* in order to account for the distinct cognitive strategies (MacCleod, Hunt, & Mathews, 1978; Mathews, Hunt, & MacCleod, 1980).

At the UAM Individual Differences Lab, we explored the supposed transformational process of the sentence into a mental diagram through the analysis of the *evoked potentials*. Cortical areas were: Fp1 and Fp2; F7 and F8; F3 and F4; W1 and W2; T5 and T6; P3 and P4; O'1 and O'2; O1 and O2. Electrical activity from five subjects was registered along with their comprehension of the sentence (Icarán, Colom, & Martín-Loeches, 1993).

The results show that when subjects are using different cognitive strategies (linguistic or pictorial), the electrical activity at the *Wernicke area* [*W*] is quantitative as well as qualitative in nature. Furthermore, in one case in which one subject was explicitly invited to use the linguistic strategy and simultaneously avoid forming any mental image, we observed that a few milliseconds before pressing the key allowing him to go on to the verification part of the task, his visual areas were highly activated (as well as the motor areas, of course), suggesting the *internal generation of a mental picture from the sentence just comprehended* (Figure 3).

Individual Differences in Working Memory Resources

The basic questions here are the following: do all the subjects experience the same degree of competition for spatial and verbal cognitive processing resources? Are the subjects more fit to hold a verbally or a spatially overdemanding task? How can working memory resources influence the organic handicaps such as usual access to visual information? Are working memory resources reorganized by virtue of the usual modes when accessing environmental information?

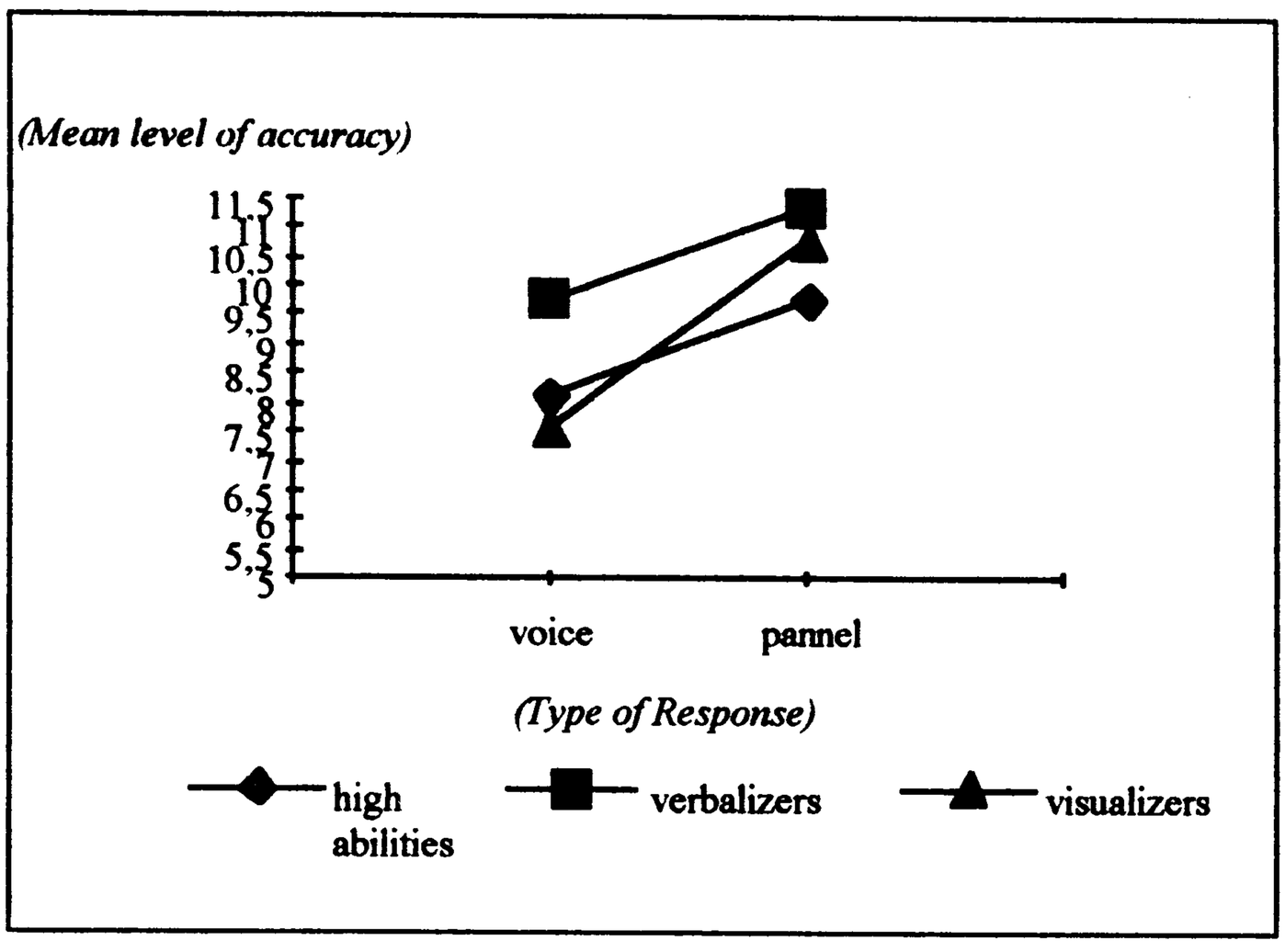


Figure 4a.

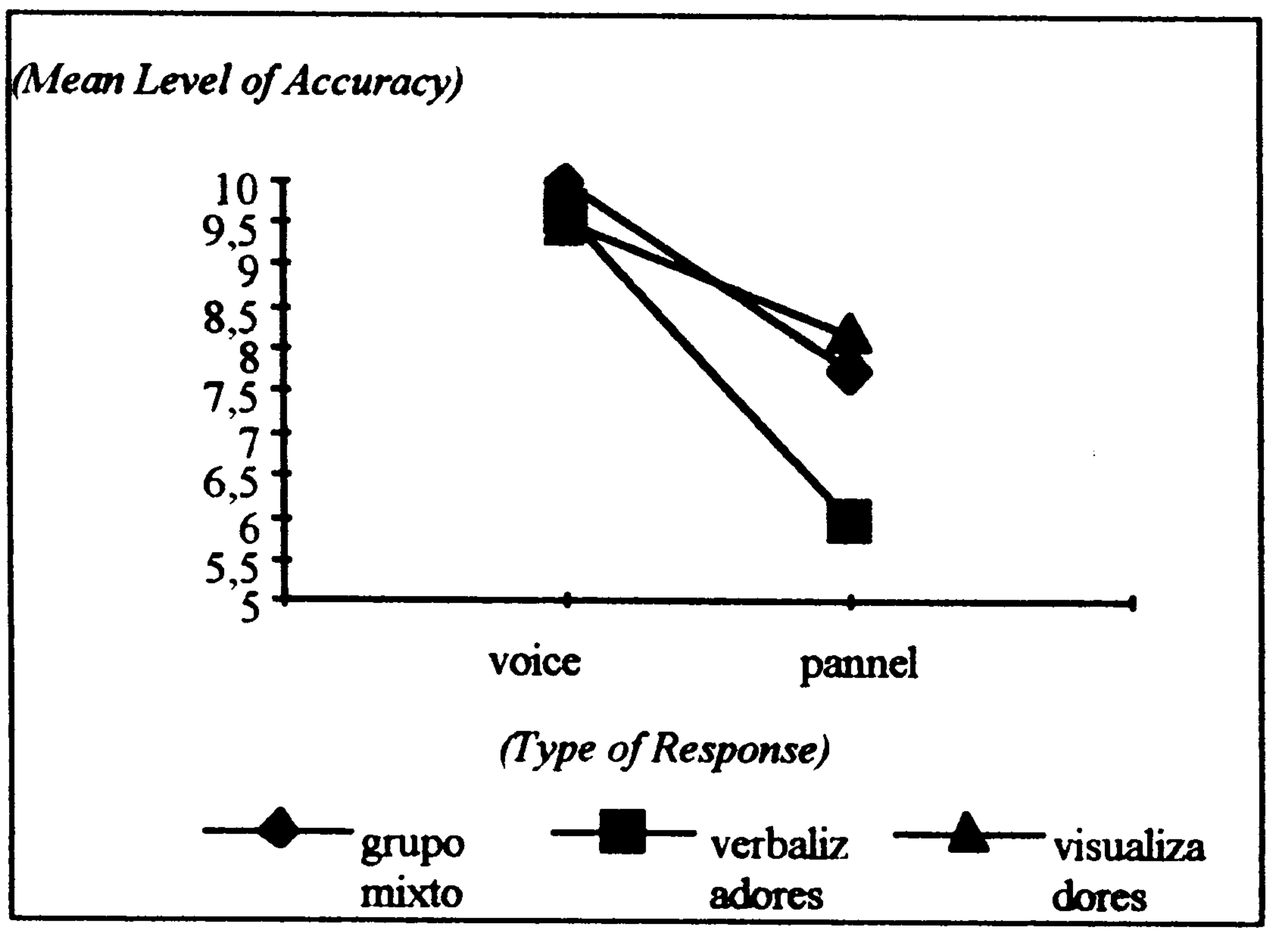


Figure 4b.

These questions where explored using the Brooks' well-known memory interference paradigm, Seamon and Gazzaninga's task researching memory retention strategies related to brain asymmetry, and the mental travel paradigm used by Stephen Kosslyn et al. (1978). The studies were conducted in our UAM Lab. and involved blind, partially blind, and sighted subjects.

Brooks' Paradigm

Here, the subject must retain a sentence or a figure of varying complexity (length or number of angles). Afterwards, they are required to perform some mental operations drawing on their verbal resources (response through voice) or their spatial resources (pointing to a panel). Subjects must therefore use the same cognitive resources to retain the information and do certain mental operations (competitive condition) or use some resources to retain information and other resources to operate and respond (distributed condition).

The variables manipulated were: type of information retained (sentence, figure) and type of response (voice, pointing to a panel). The experimental group was controlled (high abilities group, verbalizers, and visualizers; ten subjects in each group). The accuracy level was measured.

The main results are the following (Figure 4a and 4b; García, 1993; García & Colom, in press): (a) verbalizers perform better in the competing verbal condition, while visualizers perform better in the competing spatial condition, (b) these two groups are unaffected by noncompeting conditions and (c) the high abilities group is as poor as the visualizers in the competing verbal condition, but as good as the visualizers in the competing spatial condition, suggesting that subjects with high scores in both abilities are not aware of their verbal processing resources.

Seamond and Gazzaniga's Paradigm

The experimental sequence of the task is as follows: (a) a pair associate is tachistoscopically shown for 7.5 seconds (WOMAN-SEA), and the subject is asked to retain the words by repeating them mentally (in one experimental session) or by generating a mental image after the two words (in another experimental session); (b) the words disappear from the subject's visual field (blank period of 2.5 seconds) and a fixed point appears in the middle of the visual field (a plus sign, for 0.5 seconds); (c) a picture is projected to the left or to the right hemisphere during 100 milliseconds to verify the two words previously shown.

The retention strategy (verbal or spatial) and the hemispheric projection (thirty-six stimuli in each session to the left or to the right hemisphere) are manipulated, and the experimental group is controlled. Response time is the dependent measure.

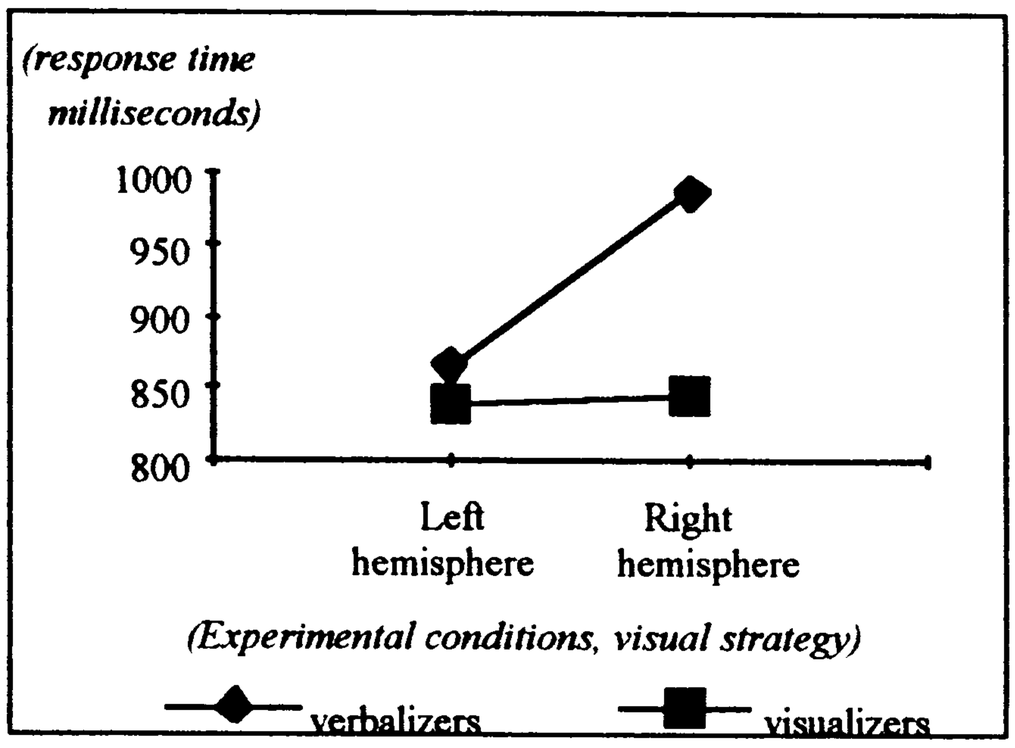


Figure 5.

The most relevant result is that when the subjects used a spatial retention strategy and the picture was projected to the right hemisphere, the verbalizers had a poor response time (Figure 4; Colom & Alonso, 1993).

This finding agrees with the classical stroop tasks previously discussed. This supports the thesis that verbalizers have a clear functional weakness when confronted by cognitive demands leading both to the non-dominant hemisphere and to their poor cognitive resources.

Mental Travel

Kosslyn et al. (1978) designed a task for studying the mental estimation of real distances between the objects of a fictitious island previously explored by haptical means. The subjects were blind, partially blind and sighted. After exploring the island, they are asked to estimate the time needed to go from any object to another on the fictitious island (Figure 6).

Perceptual status (seven blind, seven partially blind, and seven sighted subjects) is hence controlled, as well as the real distance between the objects

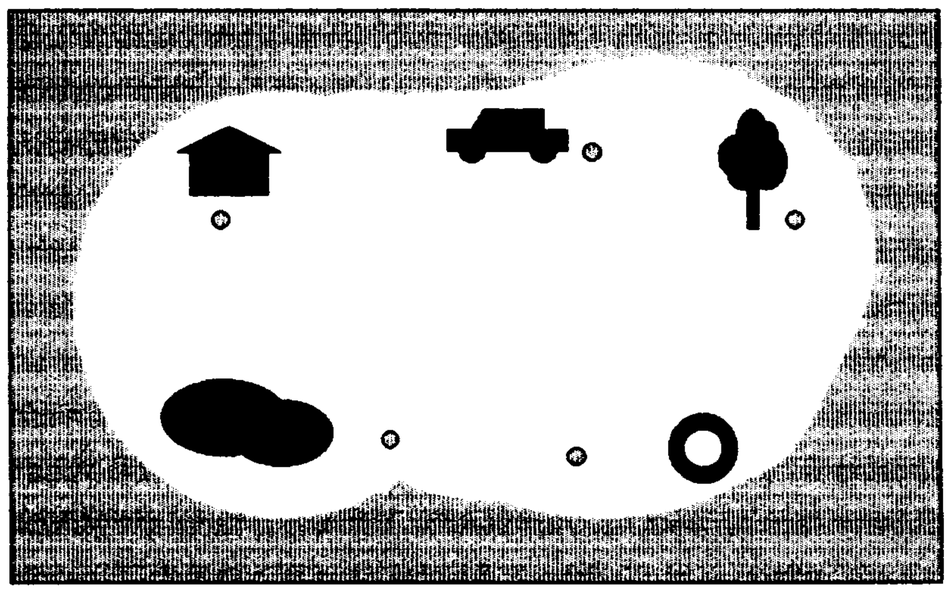


Figure 6. Kosslyn's Island Variation The Grey Dots Represent Landmarks The Five Objects Are: House, Car, Tree, Lagoon, and Pool

Table 3. Correlations Between Mental Travel Times and Real Distances

Congenital Blind	*Partially Blind*	*Sighted*
.7412**	.7318**	.5112
.7039*	.7100*	.6939*
.4559	.5488	.6836*
.7301**	.8198**	.5266
.7459**	.7156*	.4499
.6294*	.7158*	.5971*
.8303**	.4873	.7085*

Notes: Stat sig. * $p < .01$
** $p < .001$

on the fictitious island. Subjects explore and memorize the island, and then mentally estimate the time needed to travel between any pair of objects.

The findings were the following (Blanco et al., 1992; Colom et al., 1992; López et al., 1994): (a) blind subjects can estimate the distances between the objects more accurately than sighted subjects (Table 3); (b) blind subjects travelled more slowly than partially blind and sighted subjects (3175,83 msec vs. 926,73 msec and 861,1 msec.); and (c) blind subjects scanned fewer haptical patterns but in a much more organized fashion.

Individual Differences in Reasoning and Problem Solving Resources

Basic questions here were: are reasoning processes modified by virtue of content? Are structural restrictions important when reasoning with various kinds of content? Is memory span an important issue in reasoning speed and accuracy?

These questions are answered by studying deductive reasoning (linear syllogisms) and inductive reasoning tasks (analogies). The content is the manipulated variable and the formal complexity of the problems is controlled. With both types of problems (linear syllogisms and analogies), standard and non-standard paradigms were used (Colom, 1995). Furthermore, blind and sighted subjects were compared while solving linear syllogisms with different contents.

Linear Syllogisms

Three elements are related in two premises with a question about their relationships:

John is better than Pete
Pete is better than Bob
Who is best?

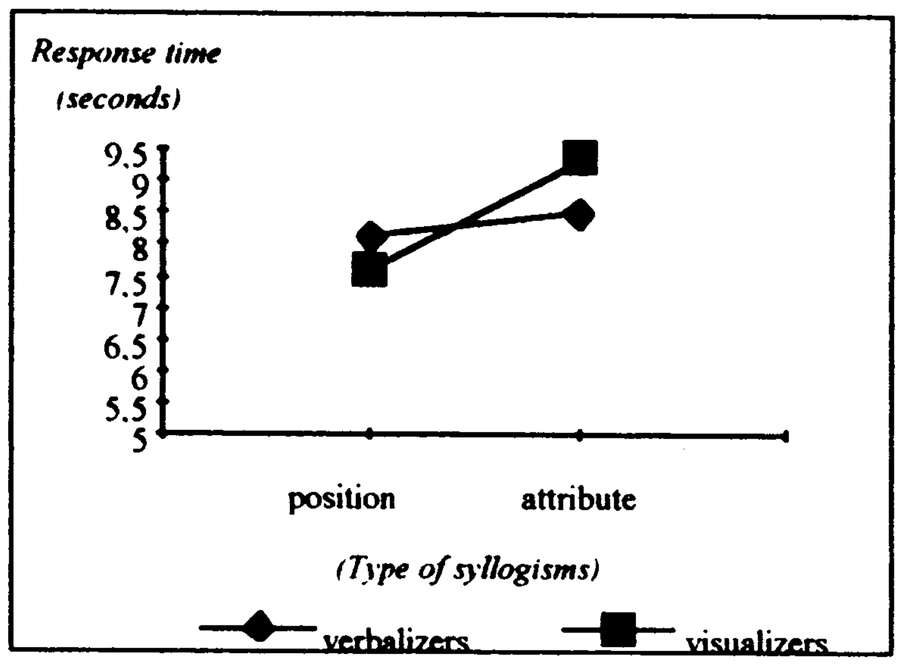


Figure 7.

Content is manipulated (positions—above, below, etc. vs. attributes—better, worse, etc.) and formal linguistic complexity is controlled (lexical marking, functional relations, etc.).

Syllogisms were presented in two standard versions: (a) 32 problems in a paper and pencil test format, and (b) 128 problems in a computerized format. A non-standard version is also employed in which problems are split into their three natural segments (first premise, second premise, and question), allowing the subject to organize the resolution sequence on his/her own (which order is preferred by the subject to organize the examination of the problem information?).

Findings about accuracy level when employing the paper and pencil test format showed that visualizers are more accurate than verbalizers in positional syllogisms, while verbalizers are more accurate than visualizers in attribute syllogisms (Colom, 1989).

In the computerized version, the results for the five subjects belonging to each experimental group were the following (Figure 7; Colom et al., 1991): (a) verbalizers have the same response time for positional and attribute syllogisms; and (b) visualizers are much slower in attribute syllogisms.

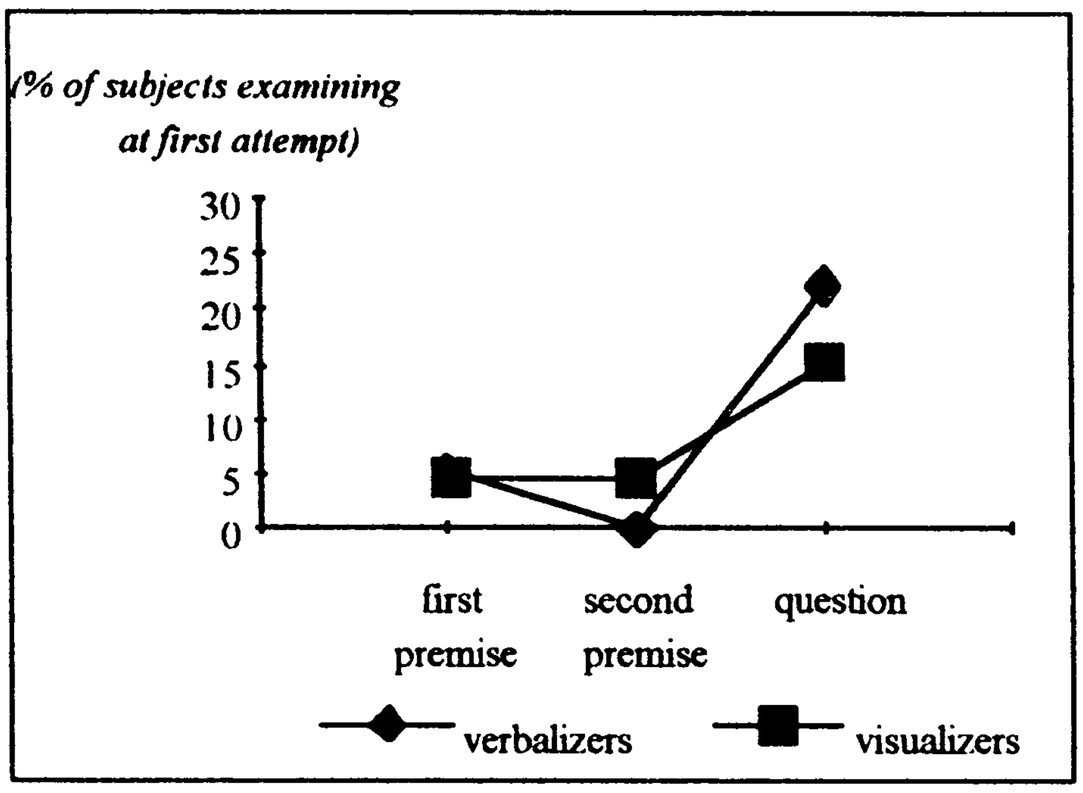


Figure 8.

These findings suggest that verbalizers process information as if comprehending any linguistic message, so the important issue is linguistic complexity and not the content of the problem, while visualizers have additional problems (i.e., need more time to respond) generating mental diagrams when the content is not easy to visualize (attributes).

In the non-standard version, the results suggest that besides the experimental group, all subjects begin by examining the question, and then the premises (Figure 8). In view of this result, it would be advisable to further explore the classical logical resolution sequence (first premise, second premise, and question) proposed in the past thirty years (Colom & Juan-Espinosa, 1990; Leo, 1993; Leo & Colom, in press).

We compared performance patterns of blind and sighted subjects engaged in solving linear syllogisms with different contents (haptical, spatial, verbal, and visual) and controlling linguistic complexity. Processing time and level of accuracy were measured, as a function of both the problem content (eighty problems, sixteen with each type of content) and the experimental group (ten subjects belonging to each experimental group).

These were the basic findings (Figure 9a and b; López, 1993; López & Colom, in press): (a) sighted subjects are faster than blind subjects, regardless of content; and (b) the blind reason with much more accuracy than sighted subjects, also regardless of content.

Therefore, blind subjects appear to be more capable of preserving the elements for reasoning for a longer period of time, a temporal advantage which improves their accuracy level. The temporal pattern of blind subjects is very diverse, and they need more time to arrive at solutions which involve certain contents (i.e., visual). However, the additional time required to solve some problems does not lead to an improved accuracy level and this finding suggests that additional time is employed to ensure a given accuracy level.

By contrast, sighted subjects show a much more regular temporal pattern, indicating that they use all the time available to them before they inevitably lose the information in working memory. This situation leads to impulsive responses, decreasing their accuracy level.

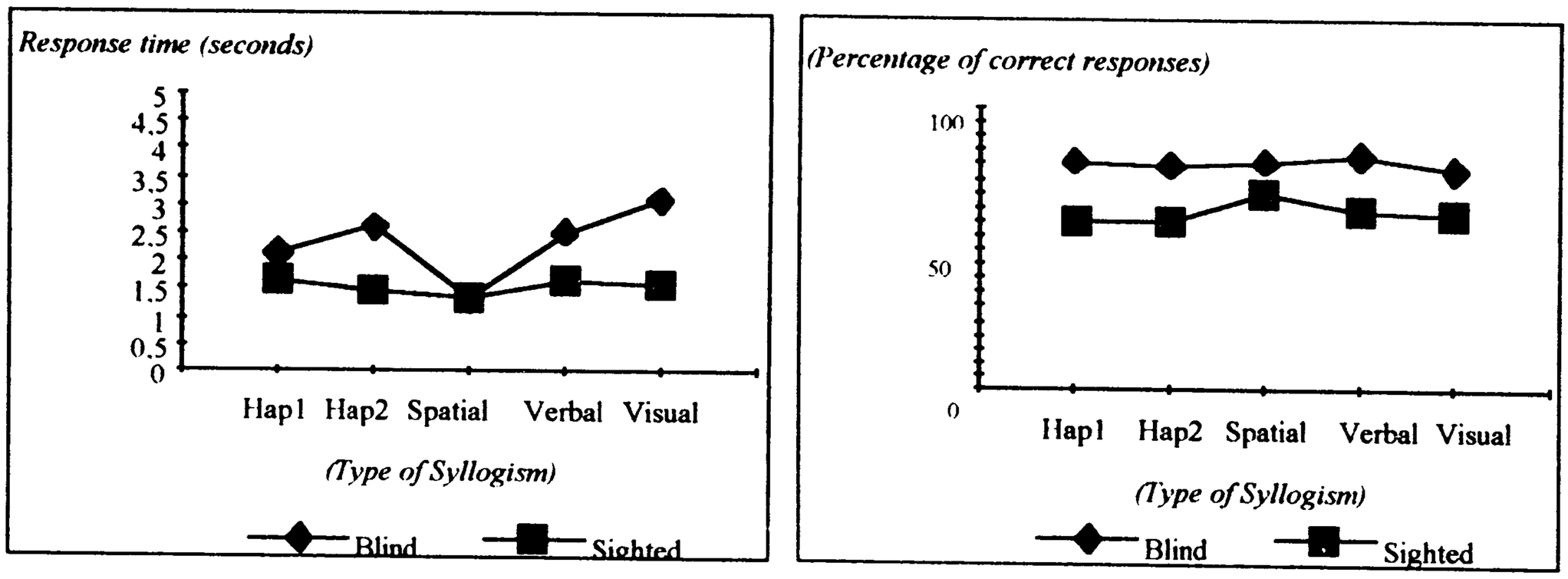


Figure 9a & b

Summorizing, blind subjects appear to be more versatile in processing information in working memory, a finding congruent with the WAIS subtest of digit span score, the blind scored higher than sighted subjects (13.6 vs.10.7).

Analogies

Verbal and figurative analogies were employed, as well as two formats to present standard (paper and pencil test format) and non-standard problems (problems are split into their three natural subparts: A is to U -domain-: as A -range- is to Z, X, Y, W -alternatives).

The results obtained with a standard version, measuring the accuracy level with thirty-two problems on a paper and pencil test format, show that the more complexity is for the figural analogies, besides the experimental group. However, when the time allowed for the verbal analogies is decreased in a second standard version, verbalizers are more accurate than visualizers (Colom, 1989; see Colom et al., 1990).

In the non-standard version, we also observe greater complexity of figural analogies, through a great number of queries needed to reach the response. Furthermore, increases in the number of queries do not lead to increased accuracy. In any case, visualizers query figural analogies with a greater intensity, compared with the fewer number of queries from verbalizers.

For the sequence of queries (domain, range, alternatives) the subjects maintain the following logical order: The person first sets out to deduce the basic rules, and then examines the alternatives, a sequence which contrasts with the results of the non-standard version using linear syllogisms (Palacios, 1993; Palacios & Colom, in press).

CONTEXTUAL MEDIATION AND LEARNING POTENTIAL

Context demands are not the only source of variance within intelligence differences, as we can see in Figure 1; context also accounts for cognitive processes and abilities differentiation. This is not only because environmental demands require certain abilities (and/or processes), but because context supports all socialization processes in which society (through family, school, social networks, etc.) *mediates* in learning.[2] It can be stated that the most accepted assumption in this field is that intellectual abilities and processes are products of person-context transactions.

How can we conceptualize contextual mediation? From a behavioral standpoint, Staats (1968, 1971, 1990) has established two sites for causation for intellectual behavior: past learning conditions and current environmental conditions. While the former site determines or mediates all intellectual basic behavioral repertoires[3] (in interaction with biological resources) during a given period of time, current conditions can *re*-mediate deficits in intelligence behavior:

> The most desirable way of eliminating spirals of deterioration (products of inappropriate original learning conditions) lies in providing the positive learning conditions the child requires to acquire his adjusted behavioral repertoires (Staats, 1971, p. 318).

From a cognitive position, Feuerstein and his colleagues have also pointed out two categories of etiological factors of intellectual deficits: distal and proximal. According to them,

> Distal factors comprise those determinants that do not lead directly and invariably to certain specific outcomes but have, rather, a trigger effect (genetics factors, organicity, poverty of stimuli, disturbance in child or parent, socio economic disadvantages, etc.). Proximal etiologies comprise those determinants that when riggered by distal factors lead directly and almost invariably to specific outcomes...when for certain reasons the distal determinant does not succeed in or is prevented from triggering the proximal one, the expected end product may not necessarily appear. The most important proximal factor is the mediated learning experience conceptualized by Feuerstein and Rand under the acronym of MLE, or a lack of it (Feuerstein, Krasilowsky, & Rand, 1974, p. 205).[4]

As we can see, the importance of contextual mediation, both in learning and in intelligent behavior, has been emphasized from different theoretical positions and, from this general standpoint, an important research program on intelligence study has been developed: the so-called "cognitive training" approach (Sternberg, 1981).

There are several antecedents of this approach in the study of intelligence and aptitudes. The Spaniard Ballesteros (1899) could be considered a forerunner of the training approach, since he developed a series of tasks and exercises in order to train (i.e., teach) mental abilities. It is also widely acknowledged that the Frenchman Binet (1909) (one of the fathers of psychometric intelligence tests) developed the concept of "mental orthopedics" to help compensate for education (Brown, 1985; Brown et al, 1992). But, as is pointed out by the most relevant authors in the field, it was Vygotsky, during the thirties, and, more recently, Feuerstein who provided the historical background to modern research in this area (Lidz, 1987b).

Following the general tendency in this field, cognitive training researchers have developed their program research based on mediated learning by looking critically at psychometric concepts and the assessment of intelligence. According to Bransford, Delclos, Vye, Burns and Hassesbring (1987), there are three important criticisms:

1. Intelligence concepts and measures do not take into consideration the responsiveness of a child to instruction or mediated learning,
2. IQ and other psychometric measurements are only products of learning, and they disregard learning processes,

3. traditional assessment procedures do not provide prescriptions regarding potentially effective intervention techniques.

That is, this approach lies between two important alternatives in the field: emphasis of modifiability vs. stability of intelligence and mental abilities, and emphasis of processes vs. products or abilities. But cognitive training researchers attempt to bypass the inadequacy of IQ, *g*, and other specific mental abilities in revealing information about important problem-solving strategies necessary to produce positive intellectual changes. Also, by assuming that children in general, and the mentally handicapped in particular, can be taught to become more competent, they emphasize cognitive strategies and other relevant emotional and motivational factors in improving intelligence. On the basis of contextual mediation or mediated learning, however, cognitive training researchers need to develop new concepts and procedures for measuring or assessing a subject's cognitive modifiability and/or cognitive processes in order to enhance remediation.

One of the most important assessment concepts in the cognitive training field is learning potential. Learning potential (LP) is a relatively new construct which is based on a "dynamic" assessment (test-training-retest) the complement the structural concept of intelligence and mental ability assessed by means of so-called "static" psychometric procedures (Haywood, Filler, Shiffman, & Chatelant, 1975).

It is generally accepted that LP is the "distance between the actual development level as determined by independent problem-solving (including psychometric intelligence measurement devices) and the level of potential development as determined through problem-solving under adult guidance or in collaboration with more capable peers" (Vygostky, 1934/1978).

But, LP could be understood as a conceptual network since, with very close definitions, several other terms have been used in this field: *zone of proximal development* (Brown & French, 1979; Campione & Brown, 1977; Vygotsky, 1934/1962); *dynamic assessment* (Feuerstein et al. 1979; Lidz, 1987b), *learning tests* (Guthke, 1982), *interactive assessment* (Haywood & Izuriel, 1992), *testing-the-limits* (Carlson & Wieldl, 1978; Kliegl, Smith, & Baltes, 1989), *mediated assessment* (Vye, Burns, Declos, & Bransford, 1987) (for a revision see Lidz, 1987a; Brown, Campione, Webber, & McGilly, 1992). Nevertheless, this network of concepts is closely and tightly related not only because of the commonly accepted definition, but also because of the basic research method used.

Learning Potential: Basic Characteristics

Test-training-posttest is the methodological paradigm by means of which LP is assessed. That is, a test is administered in its standard condition, and afterwards the subject is trained by the examiner in solving a similar (never identical) task. Finally, a posttest is again administered.

In this test-training-posttest methodological paradigm the focus can be on four main features: the task/test administered, the training itself, the subjects' observed operations, and the subject characteristics. In keeping with our model, the mediator uses certain problems and tasks to teach mental processes and assess the products of subject learning. Let us briefly review these four issues.

Assessment Tasks

Tasks commonly used in assessing learning potential come from standard (psychometric) viso-spatial (culture-free) intelligence tests. Progressive matrixes (Budoff, 1970; Fernández-Ballesteros, Calero, Campllonch, & Belchi, 1987, 1989; Fernández-Ballesteros & Calero, 1993; Feuerstein et al., 1979), Khos' blocks (Hawywood & Switzky, 1974), dots organization (Feuerstein et al., 1979), and paired-associate learning (Budoff, 1987a, 1987b) have been some of the most common materials used in LP assessment testing. Verbal material is also used, such as words (Feuerstein, Haywood, Rand, Hoffman, & Jensen, 1980) or letters series (Ferrara, Brown, & Campione, 1986) retention. It is assumed that all these selected tasks tap "higher mental process" in subjects, and that their performances can be improved by training them in underlying strategies.[5]

Since it is assumed that our society (social context) reinforces mainly verbal aptitudes (Budoff, 1979; Feuerstein et al., 1979), it is important to emphasize that task selection has been guided by "unfamiliarity" as well as the avoidance of the use of verbal tasks. The reasons for these requirements have been described by Budoff:

> The child is taught to understand how to think about solving the problems when the content of the problems may be unfamiliar and the appropriate strategies or information request may not be readily apparent.....(Also), we have demonstrated that many children with a prior history of failure in school and the usual low IQ score that accompanies such failure have the ability to work with abstract reasoning problems that are basically nonverbal in their response demands, as well as with challenging nonverbal curricula (Budoff, 1987a, p. 56).

For these reasons, in order to construct LP devices, mainly viso-spatial tasks have been selected. But, in spite of the fact that viso-spatial psychometric tasks are commonly used in the cognitive training approach, authors in the field have not taken into consideration the relevant psychometric investigation of test materials and subject resources (see earlier), as a product of the *Zeitgeist* of intelligence research.

Training

Since LP and other related concepts attempt to train test tasks, changes in the testing situation is one of the prototypical characteristics of LP assessment; that is, after standard test administration, the examiner interacts with the subject in a dynamic situation (more or less standardized) by means of which the subject could improve his/her performance and, also, provide the examiner with knowledge about the subject's problem-solving strategies or processes and other relevant behavior.

As has been pointed out by Brown, Campione, Webber and McGilly (1992)—referring to the "guided assessment" dimensions—training[6] can be classified into standardized procedures and clinical methods. In other words, in the LP investigation there are two main approaches: quantitative and qualitative ways of conducting training (Fernández-Ballesteros & Calero, 1993). The former method is systematically developed, described, and implemented (e.g., see Fernández-Ballesteros, Calero, Campllonch, & Belchi, 1987/1989). On the other hand, in clinical training a wide flexibility is assumed in the behavior of the examiner, who only follows prescriptive prerequisites (Feuerstein et al., 1979, p. 105).

In spite of the fact that clinical and standard training imply differences, the two approaches also have several characteristics in common: in both cases the examiner trains the subject in rules or mental operations needed to solve a given task/problem efficiently. Let us briefly review the conditions for reaching this goal.

In order to train efficiently, the basic training operation requires a task analysis (rational as well as empirical) of the steps, strategies, and contents needed to solve a given test/task/problem. As Sternberg (1981) pointed out:

> cognitive-training approach to understanding mental abilities can be used in conjunction with either the cognitive-correlates approach or the cognitive-components approach, or in conjunction with some other approach" (p. 1184).

Along the same lines, but from a behavioral perspective, Staats (1971) emphasized the importance of intellectual task "behavioral analysis" for training in a cognitive-language basic behavioral repertoire (e.g., see Staats, 1990). Task analysis is not only important when looking for the underlying strategies needed for problem-solving, but also when looking for the level of similarity/dissimilarity with other paired task-strategies. That is, in several LP assessments the examiner is looking, during training, for "transfer" of learning. This transfer depends on the specificity of strategies trained, which depend, in turn, on task analysis.

Finally, whatever the rationale used in selecting the basic strategies underlying a task (cognitive correlates or components analysis, behavioral analysis), training is always based on learning principles: modelling (verbal or visual), verbal instructions, verbal feedback, and verbal reinforcement are important learning components. As mentioned previously, are always present in cognitive training, learning principles ago as Staats (1971) pointed out some time. This fact has also been expressed by Feuerstein:

> Thus, the examiner constantly intervenes, makes remarks, requires and gives explanations whenever and wherever they are necessary, asks for repetition, sums of experiences, anticipates difficulties and warns the child about them" (Feuerstein et al., 1979, pp. 123-125).

Nevertheless, in spite of the fact that several authors systematically vary the various learning procedures—simple feedback, verbalization after solution, verbalization during and after solution, elaborate feedback (Carlson & Wield, 1978)—very few investigations have been conducted into penetrating both basic learning principles used in cognitive training and the balance between verbal cues and visual aids that the examiner uses in training. It seems that the *Zeitgeist* in the field does not permit communication between approaches.

Observed Operations

Finally, the LP examiner can observe the operations of several subjects during test-training-posttest sessions. From this observation, scores can be obtained concerning test and posttest, both after the performance and during training. The most common measurement indexes are the following:

1. post-pretest difference in performance, or gain score,
2. performance in posttest, or posttest score,
3. trials during training,
4. number and type of help from the examiner needed,
5. degree of transferability to other test/task/problems

These parameters, derived from the subject's behavior, allow inferences concerning learning potential and the subjects' problem-solving process failures. It is important to point out that the former types of product operations are proposed and used by the quantitative approach (Budoff, 1970; Fernández-Ballesteros, 1979), whereas the second type of operation is used by most qualitative researchers (Feuerstein et al., 1979; Karpov & Talyzina, 1986).

In any case, it should be pointed out that even if learning potential (and the other related concepts) researchers criticize IQ classifications, subjects are also commonly classified by LP scores. For example, Budoff classified subjects

based on gain score in gainer vs non-gainers.[7] Brown, Ferrara and Campione (the Illinois group; e.g., see Brown, Campione, Webber, & McGilly, 1992) classified subjects based on their efficiency in transfer: fast learners/far transferers, slow learners/narrow transferers, fast learners/narrow transferers, and slow learners/ far transferers. The level of inference increases when, based on LP scores, subjects are classified in EMR (educable mentally retarded) (e.g., Budoff, 1987a). If we look at the main concern of the cognitive training approach, the difference between IQ and LP classifications depends on the assumption that LP measurement devices—beyond assessing targets for training—also assess learning capacity/cognitive modifiability better than IQ tests do. But as in the "Russian doll," after IQ classification as well as after LP classification, subjects with low levels of intelligence/aptitudes/ability for learning, and so on, remain.

It could be summarized that, in spite of the fact that the qualitative approach emphasizes clinical methods and cognitive processes, while the quantitative standpoint proposes standardized training and quantitative products, the two approaches share several features: the tasks used and the basic verbal training conducted.

Subjects

In the cognitive training approach, subjects are mainly children and adolescents previously classified under average intelligence standard scores and, also, as EMR students. This is not the place to discuss what EMR means with respect to mental retardation classifications nor, whether this category necessarily implies the existence of an opposite one, that is, a Non-EMR category. Behind this classification is the etiology of mental retardation (see low intelligence): underlying biological (genetic causes, neurophysiological deficits, etc.) and/or environmental (social deprivation, low SES, etc.) explanations (e.g., see Feurstein et al., 1974). The main question is that children and adolescents classified as mentally retarded by IQ tests that benefit from short-term (1 hour), mainly verbal training, and who improve their performance in a (mainly) visual-spatial test, will be considered as "educable" subjects or, in other words, they can profit from the experience (e.g., Budoff, 1987a). But one question remains unanswered concerning those who do not obtain gains on LP tests. Are they biologically impaired? Are they cognitively unmodifiable?

Following our model, personal factors imply biological resources, since LP scores used to classify subjects as EMR take into consideration biological factors as a source of LP differentiation, which seem to be important. Nevertheless, as very few authors have been working in LP with "organically" mentally retarded subjects (e.g., Campione, Brown, Ferrara, Jones, & Sternberg, 1985; Roether, 1986, cited by Haywood & Izuriel, 1992), LP has been devoted mainly to "culturally-deprived" subjects.

But not only children and adolescents can be subjects for LP research, since learning potential construct is a human psychological characteristic (and perhaps a characteristic of all living organisms) and could be used throughout the life span. It is assumed that the biological and psychological aging processes parallel one another and are functionally related; in other words, that cognitive/ intellectual decline is a consequence of aging and is caused by the SNS involution. From this position, we have nothing to do with positive cognitive change in old age. The situation is quite similar to that of the children and adolescents: during old age intellectual decline has to be contemplated passively. This position has been upheld in the field of gerontology for decades. Nevertheless, recently, the cognitive training approach has also been used in the study of intelligence in old age. Several research teams are engaged in the study of "cognitive modifiability" or "plasticity" (for a revision, see Schaie, 1990) in old age, using concepts (such as "testing-the-limits") and methodological strategies borrowed from the LP concept network (e.g., Kliegl, Smith, & Baltes, 1989). Elderly people have become subjects of the research training approach in the study of intelligence.

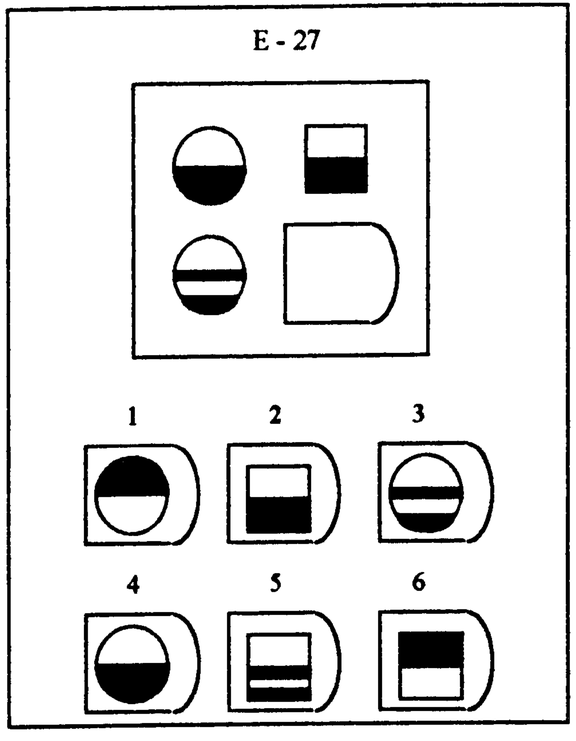
In the last fifteen years, Fernández-Ballesteros, Calero, Campllonch, Belchi and associates have been working on learning potential, from a quantitative perspective based mainly on Budoff's works (Budoff, 1970, 1975; Budoff & Corman, 1975), developing a LP assessment device called *Evaluación del Potencial de Aprendizaje* EPA ("Learning Potential Assessment"; Fernández-Ballesteros, Calero, Campllonch, & Belchi, 1987/1990) and conducting several studies with children, adolescents, and elderly people.[8] We will now summarize this research program.

Learning Potential: IQ Influence, Stability and Generalizability

After a review of the field by Fernández-Ballesteros (1979) from a quantitative perspective based on Budoff's materials,[9] a set of progressive matrixes as well as a standarized group training were adapted and/or developed for the Spanish context. Training effectiveness, influence on subjects' IQ, stability and transferability of LP scores were tested by Campllonch (198O), Campllonch and Fernández-Ballesteros (1981), Fernández-Ballesteros and Campllonch (1982), Belchi (1985) and Calero and Belchi (1989).

In our first experiment, ninety 12 to 15-year-old subjects were selected from public educational centers for culturally deprived children. Their IQs (assessed by WISC) ranged from 115 to 70. Subjects were classified by means of IQ into three groups of thirty subjects per group, with ranges of 115 to 90 (normal group), 89 to 80 (normal-low group), and 79 to 70 (low group). In order to establish experimental and control subjects, each group was divided at random into two subgroups of 15 subjects each. A similar design was used in our second experiment with forty-eight subjects, 10 to 14-year-olds, attending a school for handicapped children (63.025 IQ Mean).

EPA problem



Hunt. 1974

Gestalt Algorithm:
- continuation
- superimposition

Analytic Algorithm:
- Constancy
- Sumplement / Delete
- Expansion / contraction
- Addition / substraction
- Movement
- Composition /descomposition

This 2 x 2 Figural Analysis is solved by "constancy and supplement / delete algorithms

Sternberg, 1977

Component process:

- Encoding:
  Identify atributes and values of each term of the problem.
- Inference:
  Discover rule relating A to B
- Mapping:
  Discover rule relating A to C
- Application:
  Generate rule form D' and evaluate D
- Justification:
  Test validity of operation as performed

Examiner instructions in the EPA (Fernández-Ballesteros, Calero, Campiloch and Belchi, 1987)

- Watch carefully the picture
- First the first row, what happens there?. The circle changes into a square, and the collor?...
- What changes if we go by columns? First collomn?... Second collomn?...
- What should I draw in position D?...
- Who can tell me the number of the correct answer?...
- Of course, picture five is the correct answer. Let's see... (Transparency R-27 is showed)

Figure 10. EPA Progressive Matrix Problem, Hunt's Rational Analysis, Sternberg's Component Analysis and EPA Examiner Instructions

Ninety-six progressive matrixes were selected from previous studies by Budoff (Budoff, 1967, 1979, 1973), as basic material for training belonging to A, Ab, B, C, D, and E Raven series. The number of matrixes representing the six series are dissimilar; the C, D and E series were underrepresented in comparison to the A, Ab and B series.[10]

In our second study (Belchi, 1985), 66 subjects ranging from 12 to 15 years of age (27 males and 39 females) were selected in order to test the training specificity. The major methodical difference between our two studies was that, while in the former all experimental subjects received training in all Progressive Matrixes Series, in our second study, a search for specificity/transferability (Brown & French, 1982), the subjects were divided into three subgroups: the first group was trained in A, Ab and B problems, the second subgroup in the D series and the third subgroup in C and E.

In both experiments, the training program was administered to groups of 4-6 subjects. In order to present progressive matrix problems, a slide presentation system was employed. The presentation of each element consisted of two slides; the former showed the design problem and the second the solution.

Based on Hunt's studies (Hunt, 1974) of Raven's problems, and following Sternberg's process components for analogic reasoning tasks (Sternberg, 1977), examiner-subject interactions are systematically described and structured (See Figure 10).

In general, after presenting the slide problem, the examiner elicits an answer from the group. Correct responses are reinforced verbally at the same time the slide response is shown. From this point on, the examiner systematically tries to elicit from the group the cognitive-verbal strategies involved in reaching the correct answer. The examiner also elicits explanations for potentially incorrect solutions. Verbal and visual modeling, verbal instructions, verbal and visual feedback and verbal reinforcement are also used during training.

In test and posttest phases, the standard version of the Raven Intelligence Test (Raven, 1975) was used as a dependent variable. The Raven test was administered to all subjects twice, but while the experimental subjects received two sessions (approximately forty minutes each) of training after the pre-test, the control subjects did not. In order to assess the stability of the LP scores, two follow-ups were conducted after a one-month and a one-year interval. In order to assess construct transfer, two other general intelligence tests were administered twice (pre-and post-training) in addition to the Raven test, during the first study only: Dominos Series (TIG-1; TEA, 1976) and Thurstone's PMA Spatial subtest (TEA, 1971).

The principal results were the following:

1. The first study yielded significant differences between experimental and control subjects both in post-test and gain score ($F\,(2.168) = 119.1$; $p < .000$). The second study yielded similar results for all series.

2. Subjects with low and normal-low intelligence obtained higher gains than IQ normal subjects. Significant interactions, that is, training x IQ (F (2.168) = 3.5; $p < .03$) and Raven test x IQ x training (F(4.168) = 2.7 ; $p < .03$), were found.

3. In spite of the fact that stability in gain scores was found, both after one month and one year, differences between trained and untrained subjects were maintained after one year only in "low-intelligence" subjects.

4. Concerning the transferability or generalizability of learning, in our two studies, significant differences between the six progressive matrixes series gains were found. In our first study, the experimental subjects' gain scores in the A, Ab and B series were significantly higher than in the C, D, and E series. In our second study, when subjects were trained in a set of matrixes (A, Ab orB/ D/ or C and E), gains were found only in the trained series, that is, the results yielded no transfer from difficult (more complex) to easier series. Along the same lines, no differences were found between trained and untrained and pre-posttest experimental subjects on other general intelligence test scores (TIG-1 and PMA Spatial subtest scores); that is, no transfer construct was found.

In summary, from the first step of our LP research program we learned the following:

1. Comparisons between experimental/control groups and pre/posttest gain scores allow us to infer that our training program teaches the strategies underlying progressive matrixes. Raven posttest scores and gain scores were sustained in the follow-up, one year later. That is, our LP training was successful and our LP measures were stable over time.

2. The result that normal subjects improve less than low and normal-low subjects seems to be due to a "ceiling effect" in the A, Ab, and B series in the Raven test. In the Raven pretest, normal subjects scored close to the maximum. Even if gain percentage scores were used (instead of gain or posttest scores) normal subjects could not improve their performance on the Raven test.

3. Strategies in which subjects were trained in both studies were extremely specific. When experimental subjects significantly improved their performance on the Raven test, it was clear that training is focused mainly on strategies present in A, Ab and B matrix problems. There was no transferability, even from the most difficult/complex series (series C or D) to the easier series (A or Ab). Also, gains on the Raven test after training do not affect other "*g*" test performances. From these results, the specificity of the strategy trained and the lack of transferability of learning can be inferred.

Further research should be conducted to learn more about the generalizability/ transferability of cognitive strategies. Needless to say, based on these results, EPA (our LP measurement device) was modified, including additional of progressive matrix designs corresponding to series C, D, and E in order to have a similar number of training materials in all Raven series (Fernández-Ballesteros, Calero, Campllonch, & Belchi, 1990).

LP as Predictor of Cognitive Change in Culturally Deprived and Biologically Damaged Subjects

One of the most important goals of learning potential and related concepts is to replace IQ, and other aptitude tests, to improve predictions of learning. According to Budoff (1987a):

> The intent of learning potential assessment is to obtain an estimate of general ability derived from reasoning problems of suitable difficulty, which the child has had an opportunity to learn how to solve, and which permit a comparison with the low scholastic aptitude score (e.g., Binet or WISC Verbal IQ). If the child can demonstrate, following a short period of training on a *nonverbal* [11] reasoning task, that he or she can perform at a level approximating average peer performance, we interpret this to indicate that he or she is not mentally retarded. This reflects our definition of intelligence as the ability to learn from the experience (p. 55).

But, if after a short treatment subjects improved their performance in a intelligence test, this proves neither that these results predict improvement in a broad range of variables (as suggested by Budoff, 1987a, p. 76) when subjects receive long-term cognitive training, nor that LP scores are better predictors than IQ scores in learning situations.

Also, as we emphasized before, based on this research program several authors have investigated the value of LP in distinguishing between "culturally deprived" and other MR which can be assumed to be biological conditions.

There is contradictory evidence concerning LP differences between several types of "biological impairment or dysfunction" and culturally deprived subjects (e.g., Haywood, 1969; Diaz, 1982). In our UAM Laboratory, Belchi (1985) conducted a pilot study exploring LP among "organically" mentally retarded subjects ($N = 19$; aged 12 to 15) using gain scores from our LP measurement device (EPA). The only difference between the organic group and other mentally retarded subjects was not the ability to learn but the magnitude of learning (gains after training in the organic group were lower than in the results of previous experiments involving the non-organic group), as well as the lack of temporal gains stability. However, results from this pilot study should be extended.

In this second set of studies developed by Calero (1986), Fernández-Ballesteros and Calero (1992, 1993), our main goal was to verify whether gains in our LP instrument (EPA: Evaluación del Potencial de Aprendizaje) adequately predict intellectual changes after a long-term cognitive intervention program involving both culturally deprived and biologically impaired subjects. In other words, the general hypothesis is that "gainer subjects" (assessed by EPA) will significantly increase their IQs as measured by the Wechsler Intelligence Scale for Children (WISC) over the "non-gainers" after participating in a one-year cognitive program. Differences between culturally deprived and biologically impaired subjects were also expected.

The study involved forty-eight subjects (18 females and 30 males), 10 to 14 years old. Their IQ (assessed by WISC) ranged from 50 to 85 (Mean: 63.025). Twenty-four of the 48 subjects were diagnosed by a Clinical Service of Neurology as "organically impaired," while no brain damaged or biological dysfunctions had been detected in the remaining 24, assuming that these subjects were "culturally deprived." Subjects from each category were randomly assigned to the experimental and control groups.

In order to assess learning potential, the EPA ("Evaluación del Potencial de Aprendizaje"; Fernández-Ballesteros, Calero, Campllonch, & Belchi, 1987/1990) was used. The Raven Intelligence Test-General Scale (Raven, 1968) was administered before and after EPA training. As a long-term cognitive intervention program, Feuerstein's FIE was selected, since there was ample evidence of its effectiveness (Feuerstein, Rand, Hoffman, & Miller, 1980). The FIE modules administered were: Dots organization, Spatial Orientation I and III, Comparisons, Categorizations and Analytic Perceptions. These six modules of the FIE program were administered weekly, in forty-five minute sessions, throughout a school year (8 months) to the experimental groups. In order to assess the intellectual change produced by FIE as predicted by EPA, the Wechsler Intelligence Scale for Children (WISC; Wechsler, 1974) was administered before and after the FIE program. Control subjects followed the regular school program and they were assessed using Raven (before and after EPA training) and the WISC twice (before and after training), with the same schedule applying for the experimental group (for more information about procedures see Fernández-Ballesteros & Calero, 1993).

The principal results were the following:

1. The long-term cognitive training program had an effect in the expected direction, producing significant changes in the experimental subjects compared to the control ones in the WISC Verbal scores, both raw ($F = 21.154$, $p < .0001$) and standard scores ($F = 18.51$, $p < .0001$). No significant differences were observed after intervention in WISC performance scores.

2. Covariance analysis yielded a positive interaction between long-term intervention and EPA gain scores ($F = 25.133$, $p < .0001$). EPA gainers (compared with non-gainers) showed significant improvements in their intelligence scores both in experimental and control subjects.

3. Regarding the subjects' etiology (organic vs non-organic), organic subjects, in comparison with non-organic subjects, showed significant differences in their EPA gain scores (3.29 and 6.04, respectively).

4. The predicted influence of EPA gain scores on the WISC gains after a long-term cognitive intervention was significant. The post-WISC scores were better than the pre-WISC scores. Biserial correlation yielded a high relationship between EPA gain scores and WISC verbal gain scores ($r = .71$, $p < .001$). While the regression coefficient of raw WISC verbal scores before and after long-term cognitive intervention was not significant ($r = .086$), the regression

coefficient between EPA gains and Verbal raw WISC gain score (after treatment) was .4151 ($p < .01$). Finally, ANOVAs conducted with high and lower pre-WISC total scores did not yield any significant differences between control and experimental subjects.

5. Regression analysis showed, at the same time as EPA gains in experimental groups, their stability after one year. "Organic" subjects showed a lower level of stability in the one-year follow-up.

In summary, from the first step of our LP research program we learned the following:

1. Our LP measurement instrument (EPA) scores were better at predicting positive changes in the WISC verbal score after a long-term cognitive intervention program than WISC scores. EPA scores were better predictors of improvement in verbal IQ scores even in the control group. This is an important result, since EPA could predict not only the effects of the FIE program but also the benefits of other "natural treatments" (education, family, social context always imply non-systematic "interventions"); nevertheless, this is a post-hoc explanation that should be tested.

2. "Organic" subjects obtained lower LP scores as well as fewer benefits from our long-term cognitive intervention program. Nevertheless, the number of subjects involved in this research, as well as our results, are insufficient to conclude that organically mentally retarded subjects are non-educable or modifiable subjects; to a lesser degree than the others, they can obtain benefits from cognitive interventions.

3. The EPA gain scores were stable after the one-year follow-up in the total sample. Nevertheless, non-organic or culturally-deprived EPA gain scores were significantly more stable than those achieved by organic subjects.

An important question arose from these results concerning the type of intelligence improvement achieved: after a long-term cognitive intervention program, subjects improved their verbal scores significantly (they did not improve significantly on either total or performance scores). Our question is: how is it possible that changes in WISC verbal scores[12] produced by a long-term cognitive program were more accurately predicted by performance on the EPA, whose material is visual-spatial, than by the WISC verbal scores? This research question also concerns the nature of other LP measurement procedures and the author's predictions concerning LP and IQ tests. For example, Budoff (1987a) stated:

> The hypothesis underlying potential assessment that verbal skills are independent of, and that poor verbal skills may mask, general ability to reason was supported by low partial correlation coefficients involving WISC Verbal IQ (p. 61).

Our data do not support this and other statements concerning the non-verbal nature of LP measures. It should be mentioned that the basic EPA material involves progressive matrixes. Beyond the fact that progressive matrixes measure inductive reasoning (Carpenter, Just, & Shell, 1990) and beyond the operations (Hunt, 1974) and components (Sternberg, 1977a) needed to solve them, we have evidence—from a psychometric as well as a neuropsychological perspective—of their visual-spatial nature (e.g., Carpenter et al., 1990; Risberg & León-Carrion, 1988). On the other hand, as mentioned before, LP training is, essentially, verbal; according to Campione and Brown (1979), our results are congruent with the hypothesis that with LP assessment executive control, verbal operations can be trained. That is, LP measures, using mainly verbal training to teach visual-spatial tasks, can train verbal operations, rule-governed behavior, or metacognitions (Catania, Matheus, & Shimoff, 1990; Luciano, 1992; Sternberg, 1977b). From this point of view, LP scores—obtained by training through metacomponents of visual-spatial problems—predict Verbal IQ gains after a long-term cognitive treatment—which train through metacognitive components of visual material. Although this is a post-hoc explanation of our results, the main point is that we have not found any research which investigates the importance of this cross-task training effect between spatial material with mainly verbal training in LP measurement devices.

If we look at other research programs summarized in this chapter devoted to the interactions between personal resources and context demands, we can make new predictions clarifying these results.

A first attempt to predict these effects has resulted in a new correlational analysis of our data. Our prediction is as follows: "verbal" subjects will improve their performance in our EPA training program better than "visual-spatial" subjects. In order to test this prediction, our research subjects were divided into four groups by means of verbal (Vocabulary) and performance (Blocks) subtest scores from the Wechsler Intelligence Scale for Children (WISC) test. The four categories were: verbal (high raw score in Vocabulary and low raw score in Blocks), visual-spatial (high raw score in Blocks and low raw score in Vocabulary), high verbal-visual (high raw scores, in both Vocabulary and Blocks) and low verbal-visual (low raw scores in both Vocabulary and Blocks). Our prediction is that high verbal and high verbal-visual subjects obtain better EPA gain scores than high visual-spatial and low verbal-visual subjects. As a first confirmation of this prediction, the data presented in Table 3 show that high verbal and high verbal and visual-spatial subjects and high visual-spatial subjects displayed greater improvements in their performance on the Raven test after EPA training than high visual-spatial subjects or low verbal-visual subjects did .

Since this is a post-hoc test, new research should be conducted in order to assess the effect of resource (abilities) interactions with task demands in potential learning assessment and contextual mediation.

Table 4. Mean, SD, and ANOVA of EPA Gain Scores of "Verbal" and "Spatial" Subjects

		SPATIAL SUBJECTS	
	HIGH	*LOW*	*TOTAL*
VERBAL SUBJECTS			
HIGH	Mean= 7.583	7.154	7.360
	SD= 4.944	6.998	4.356
LOW	Mean= 4.222	2.857	3.391
	SD= 5.932	3.110	4.356
TOTAL	Mean= 4.926	6.143	$F(1/46)= .56$
	SD= 5.677	5.515	$p < .46$
		$F(1/46) = 6.81$	
		$p < .013$	
Interaction: $F(1/44)= 0.09$; $p < .77$			

Learning Potential in Old Age

The LP approach has evolved in a developmental and educational context; since the main goal has been how to assess children's cognitive modifiability or educability, research programs in LP have been developed working with subjects ranging from kindergarten to high school age. However, all criticisms levelled at the psychometric (so called static) approach to intelligence can also be applied to other age groups, even to the elderly. As emphasized by Schaie (1990), referring to Gould (1981), biological assumptions from a Darwinian tradition cannot be applied to human cognitive development across age and aging. From a socio-historical point of view, human subjects receive their cognitive heritage from previous generations, but unlike biological characteristics, which cannot be learned, cognitive resources can be "inherited" from the context. Although biological principles govern the brain and its development and evolution, intelligence is governed by socio-historical standards (Labouvie-Vief, 1984). From this point of view, plasticity or modifiability would be a human characteristic throughout the life span; that is, in our terms, contextual mediation is also possible in old age.

In recent years, modifiability or plasticity have become the subject of psychogerontology research (e.g., Baltes & Willis, 1982). From the point of view of the person and her/his resources, testing-the-limit, (e.g., Kliegl & Smith, 1989) and reserve capacity (e.g., Baltes, Dittmann-Kholi, & Kliegl, 1986) are new assessment concepts in the field of research on aging and cognition. Experimental research results in the field make clear that older subjects can benefit from contextual mediation or cognitive training; even cognitive intervention can compensate cognitive declines caused by aging

(e.g., Schaie, 1990). These results suggest the continuity of the modifiability of intellectual performance through contextual mediation (cognitive training) across the life span.

After reviewing the literature, several points should be emphasized:

1. Subjects who participate in cognitive training studies are in good health, are "young" elderly persons, and comparatively, well educated (e.g., Willis, 1987).
2. Reserve capacity or intellectual modifiability through cognitive training, in old age always take into consideration the effect of long-term cognitive training, and very few studies use short-term training as applied in learning potential studies.
3. According to the remarks made in previous studies, training results in short-term transferability (e.g., Willis, Blieszner, & Baltes, 1981). As in other stages of life, learning potential or intellectual training effectiveness should have an effect not only on closed test performances, but also on practical intelligence (Schaie 1990, p. 300).

Two main questions guide our research in the field of aging (Fernández-Ballesteros & Calero; 1992, 1995): (1) Is EPA short-term training better able to change Raven Test scores in elderly poorly-educated subjects when compared with other brief training programs related to everyday competence (spatial and social skills)? and (2) If short-term training yields positive results, how much transference will it have?.

Ninety elderly (36 women, 54 men) voluntary subjects (age range: 59-87; Mean age: 67.87) participated in this study. Sixty-four (20 women, 44 men) lived at home and attended Senior Citizen Clubs. Twenty-six subjects (16 women and 10 men) lived in retirement homes for autonomous elderly individuals. More than 93 percent had less than 4 years of elementary education but all were literate (selected from a group of volunteers who were administered a reading and writing test). ANOVA analysis yielded no initial differences in any of the socio-demographic variables between those subjects living at home and those in residential centers.

The study was based on an experimental-control group design with random assignment to four conditions. Four steps were followed: pretest, cognitive training (three training conditions vs. placebo control group), posttest and follow-up (after three months of the training). Only 62 subjects participated in the follow-up. In the pretest, posttest and follow-up, a battery of tests was administered; the order of test administration was random. Each training group participated in one of the three training conditions: EPA training (or Inductive Reasoning with Progressive Matrixes training) (Fernández-Ballesteros, Calero, & Belchi, 1990) lasted for a total of five one-hour sessions; Spatial Practical Orientation (SPO) training (based on Feuertein's Spatial-practical orientation

module FIE; Feuerstein et al., 1980) lasted for ten one-hour sessions; and Everyday Problem Solving (EPS) training based on De Bono's PDP (De Bono, 1976) lasted for ten one-hour sessions.

Six tests were administered: a domain test and a transfer test for each training. For EPA training, the Raven Intelligence Test (RPMt), General Series (Raven, 1968) was used as a domain test, and the G-Cattell test (CGt: 1,2, and 4 subtests) (Cattell, 1975) as a transfer test. For Spatial Practical Orientation (SPO) training, the Practical Orientation task (PSOt) was used as a domain test (constructed for this research) and the Yuste Spatial Orientation task (YSOt) (Yuste, 1983) as a transfer test. Finally, for Everyday Problem Solving training (EPS), the domain test was borrowed from the PDPt (Sánchez el al. 1983), and the transfer test from the Spanish adaptation of the Everyday Problem Solving Inventory (EPSI) by Cornelius and Caspi (1987).

The principal results were the following:

1. In order to establish the configural test interrelationship, that is, an empirical transfer gradient between tests, correlational analysis were conducted. Based on pretest correlations, an empirical transfer gradient, or hierarchy of tests according to the proximity of training, was found for the total sample ($N = 90$). The Raven test correlated significantly with the G-Cattell test (.66, $p < .01$), the Practical Orientation task (.54, $p < .01$), the Yuste Spatial Orientation test (.52, $p < .01$), and the PDP test (.35, $p < .05$). The G-Cattell test correlated significantly with the Practical Orientation task (.53, $p < .01$), and the Yuste Spatial Orientation test (.48, $p < .01$). Finally, the Practical Orientation task correlated significantly with the Yuste Spatial Orientation test (.49, $p < .01$). This so-called transfer gradient was maintained in the posttest and follow-up data.

2. Concerning training effects, after EPA training, the Raven (domain) test gain score was 9.24, an increase of one Standard Deviation ($F = 12.6$, $p < .0001$), and the score for the G-Cattell (transfer) test was 2.71, an increase of 1/5 Standard Deviation ($F = 3.23$, $p < .026$). After Practical Spatial Orientation (PSO), training effects in the domain (PSO; $F = 6.48$, $p < .0005$) and transfer (YOSt; $F = 2.55$, $p < .06$) tests were found. PSO training also yielded transfer results in the Raven test, with trained subjects gaining a pre-post test 1/5 Standard Deviation. No significant effects were found for the third training program administered. At the same time, the control group yielded no significant pre-post-test differences.

3. Follow-up results analysis showed that the effect on those abilities improved as training persisted after three months.

In summary, concerning LP and transferability in old age, we learned the following:

1. Consistent with other studies on modifiability and plasticity among highly-educated elderly persons, we found that poorly-educated elderly people improve their scores on the Raven test after our EPA short-term training (5

hours). This training resulted in transfer to the G-Cattell test. The Practical Spatial Orientation ten-hour training also had an effect in both the domain and transfer test. Its transferability was the broadest, improving subjects' performance significantly on the G-Cattell test. As in other studies, stability (after three months) was found for both effective trainings. Everyday Problem Solving did not produce effects in any of our outcome measures, and neither have we found correlations between them. We cannot reach any conclusion concerning which part of the variance can be attributed to the training program's lack of efficacy or to the lack of validity of the measures used to assess social competence.

2. It seems that the training transferability gradient can improve the use of empirical data to support predictions about transfer in learning potential assessment. We have obtained better transfer results than in previous experiments.

After this study we learned that the LP paradigm is useful for investigating contextual mediation during old age. Trained subjects improved their intellectual abilities, but what about competence in everyday life? Since our purpose to improve is not only intelligence scores but also practical intelligence and/or everyday problem-solving, new LP devices using everyday adaptive behaviors from everyday cognitive problems are needed, as well as effective training. More LP creative procedures should be developed and tested in the near future.

CONCLUDING REMARKS

In summary, as far as the person-context adjustment is concerned, our results lead us to draw the following conclusions.

In the first place, people appear to have organized general concepts of intelligent behavior into prototypes, but the general organization differs according to context. This difference could be interpreted as a result of the acculturation process, as opposed to the traditional culture which prevails in the rural context. Despite this difference, people from our two contexts seem to share a general prototype of the intelligent person.

Second, the presence or absence of schools, urbanization, and so on—the indicators of acculturation—influence both the structure and contents of intelligence prototypes. Evidence of influence on structures can be found in: (a) the order of importance of abilities, and (b) the presence of specific culture-bound factors.

Taken in general, our results with the Fang people are consistent with the law of cultural differentiation, and can be taken as an example of the acculturation mechanism proposed by Berry and of its influence on people's beliefs about intelligence.

As far as the personal resources and contextual demands are concerned, our concluding remarks are:

1. Mental tests allow a sampling of the subjects' cognitive resources, but the resources sampled are not enough to predict specific human performance in some cognitive tasks with any precision.
2. The high abilities group hence shows performance patterns that point to ineffective processing of verbal information in working memory.
3. As is suggested by the performance patterns observed in blind subjects, *memory span* appears to be an important issue in solving linear syllogisms, that is, in complex human problem-solving. This finding leads to the next question: will performance improve by training retention strategies in working memory so that subjects can reason more comfortably? (see Carpenter et al., 1990; Kyllonen, in press; Kyllonen & Christal, 1990).
4. Persons are *naive economists*. In deductive reasoning tasks, they begin by examining what is questioned in the problem, and then scan the information within the two premises. In contrast, when they solve analogies, the situation changes in many respects: the piece of information involving alternatives is the last one examined, so that working memory does not overload. We think *these data must be explicitly considered when training human cognitive performance* (see also Kyllonen, in press). Therefore, the aim is to design *non-invasive training systems* (Snow & Swanson, 1992).
5. Not all subjects have all the necessary cognitive resources to solve all the tasks designed by (depraved!) psychologists. For example, the findings show that subjects do not react equally to verbal and spatial processing resources overloads. Again, this could be an important issue in designing training systems. *Psychologists must be aware of persons' particular resources for adapting training systems* (Kyllonen, in press).

Finally, context can mediate appropriately in learning and in the improvement of subjects' intellectual performance, both in children and adolescents and in old age. Learning potential measures are better indicators than IQ scores for predicting intellectual performance and learning in long-term cognitive intervention programs. Also, "organically" or biologically impaired children showed learning potential, transfer and stability to a lesser degree than "non-organic" children. The elderly can benefit from the effects of short-term training; plasticity seems to be a condition throughout the life span. It is also important to emphasize that we need to learn more about the generalizability and transferability of our LP training. We must pay more attention to both individual differences and task/problem differences in our LP measures. Finally, longitudinal studies are necessary, involving children

and adolescents and old people, both to test predictions from LP measures and to learn more about the long-term effect of mediated learning.

NOTES

1. Psychodiagnostic, Gernontology, and Individual Differences Laboratories.
2. Even data from animal research support the importance of context (e.g., stimuli deprivation, contect enrichment) in physical and neurophysiological organism's growth and decline.
3. Three categories of the intelligence basic behavioral repertoires (or BBRs) are proposed by Staats: cognitive-language, sensory-motor, and emotional-motivational, which are learn through a cumulative hierarchical acquisition by learning principles.
4. It cannot be said that Staats' past environmental conditions are equivalent to Feuerstein's distal factors, but they have parallel conceptual developments.
5. It is important to emphasize that *coaching* is a related concept with antecedents throughout the history of psychometric testing (see Anastasi, 1981).
6. Brown et al. (1992) called this dimension in the field of learning potential (or "guided assessment") "interaction."
7. Budoff (1987a) has changed this classification into a continuum of learning potential.
8. As has already been mentioned, these researchers have been conducted in our Autonoma University laboratories.
9. Our acknowledgement to Dr. Budoff, once more.
10. At this moment we assume (as Budoff did) the cognitive strategies transfer. In the EPA Second Edition (1990) the number of design series are similar.
11. Underlines are added to reemphasize the nonverbal nature of LP tasks.
12. It is important to realize that the modules selected from FIE require mainly visual-spatial material, but that training is mainly verbal. That is, as mentioned before, cognitive strategies used to solve spatial material are trained essentially by means of language (verbal instruction, feedback, reinforcement).

REFERENCES

Anastasi, A. (1981). Coaching, test sophistication, and developed abilities. *American Psychologists, 36*, 1094-1102

Baker, R. R. (1981). *Human navigation and the sixth sense*. London: Hodder and Stoughton.

Baltes, P.B., Kliegl, R., & Dittman-Kohli, F. (1988). On the locus in research on the plasticity of fluid intelligence in old age, *Journal of Educational Psychology, 80,* 3, 392-400.

Baltes, P.B., & Willis, S.L. (1982). Plasticity and enhancement of intellectual functioning in old age. In F.I.M. Craik & S. Trehub (Eds.), *Aging and cognitive processes.* New York: Plenum Press.

Ballesteros, F. (1899). *Educación, Didáctica Pedagógica y Practica de la Ensenanza* [Education, Didactic and Teaching Practice]. Córdoba: La Región Andaluza

Belchi, J. (1985). Potencial de aprendizaje: una crítica [Learning potential: A critic]. Universidad Autónoma de Madrid. Unpublished Dissertation.

Berry, W.R. (1987). The comparative study of cognitive abilities. In S.H. Irvine & S.E. Newstead (Eds.), *Intelligence and cognition: contemporary frames of reference.* Dordrech: Nijhoff.

Binet, A. (1909). *Les idées modernes sur les infants.* Paris: Ernest Flammarion.

Blanco, M.A. (1994). *Diferencias individuales en la tarea de stroop: Demandas vs. recursos de procesamiento* [Individual differences in the Stroop task: requirements vs. Processing resources]. Universidad Autónoma de Madrid. Unpublished Dissertation.

Blanco, F., Colom Maranón, B. R., López, R., & Leo, E. (1992). *Ceguera y desplazamiento mental en una isla ficticia: del ojo de la mente a la mano de la mente* [Blindness and mental travel withing a fictious island: from eye's mind to hand's mind]. IV Seminario Interuniversitario de Ciencia Cognitiva. Valencia, 16-18 de Enero.

Blanco, M.A., & Colom Maranón, B. R. (In press). *Visualizadores y verbalizadores en la tarea de stroop: Diferencias individuales en los procesos atencionales* [Visualizers and verbalizes in the Stroop task: Individual differences in attentional processes]. Revista de Psicología General y Aplicada.

Borich, G.D., & Bauman, P.M. (1972). Convergent and discriminant validation of the French and Guildford-Zimmerman spatial orientation and spatial visualization factors. *Proceedings of the Annual Convention of the American Psychologist Association*, Vol. 7, Pt. 1, pp. 9-10.

Bransford, J., Delclos, V., Vye, N., Burns, H., & Hassesbring, T. (1987). State of the art and future Directions. In C.S. Lidz (Ed.), *Dynamic Assessment*. New York: Guilford Press.

Brown, A.L. (1985). Mental orthopedicts: A conversation with Alfred Binet. In S. Chipman, J. Segal & R. Glaser (Eds.), *Thinking and learning skills: Current research and open questions* (pp. 319-337). Hillsdale, NJ: Erlbaum.

Brown, A.L., & French, L.A. (1979). The zone of potential development. In R.L.Sternberg & D.K. Detterman (Eds.), *Human Intelligence* Norwood, NJ: Ablex Publishing.

Brown, A.L., & Ferrara, R.A. (1985). Diagnosing zones of proximal development. in J. Wetsch (Ed.), *Culture, communication and cognition: Vygotskian perspectives* (pp. 273-305). Cambridge MA: Cambridge University Press.

Brown, A.L., Campione, J.C., Webber, L.S., & McGilly, K. (1992). Interactive learning environments: A new look at assessment and instruction. In B.R. Gifford & M.C. O'Connor (Eds.), *Changing assessments. Alternative views of aptitude, achievement and instruction.* Boston: Kluwer Academic.

Budoff, M (1987a). The validity of learning potential assessment. In C. S. Lidz (Ed.), *dynamic assessment: An international approach to evaluating learning potential* (pp. 52-81). New York: Guildford Press.

Budoff, M. (1967). Learning potential among institucionalized young adult retardates. *American Journal of Deficiency, 72,* 3, 4040-4411.

Budoff, M. (1970). Learning potential: assessing ability to reasoning the educable mentally retarded. *Acta Paedopsichiatrica, 37,* 293-309.

Budoff, M. (1973). Learning potential and educability among the educable mentally retarded. In *Studies in learning potential* (Vol. 72). Cambridge MA: RIEP.

Budoff, M. (1987b). Measures for assessing learning potential. In C.S. Lidz (Ed.), *Dynamic assessment: An international approach to evalating learning potential* (pp. 173-195). New York: Guildford Press.

Budoff, M.,& Corman, L. (1975). Effectiveness of learning potential traiing on reduction of error on Raven's progressive matrices. In *Studies in learning potential* (Vol. 48). Cambridge, MA: RIEP.

Calero, M.D. (1986a). Un estudio de validez de un test de potencial de aprendizaje [Validity study of a learning potential test]. Doctoral dissertation. Servicio de Publicaciones. Universidad de Granada.

Calero, M.D. (1986b). Cognitive training effects in an EMR adolescent group with organic etiology. Paper presented in the 21th International Congress of Applied Psychology. Jerusalem.

Calero, M.D., & Belchi, J. (1989). La generalización del entrenamiento en las técnicas de evaluación del potencial de aprendizaje [Transfer and learning potential assessment techniques]. *Estudios de Psicología, 38*, 69-81.

Campione, J.C., & Brown, A.L. (1977). Memory and metamemory development in educable retarded children. In R.V. Kail, Jr., & J.W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 367-406). Hillsdale, NJ: Erlbaum.

Campione, J.C., & Brown, A.L. (1978). Toward a theory of intelligence: contributions from research with retarded children. *Intelligence, 2,* 279-304.

Campione, J.C., & Brown, A.L. (1979). *Human intelligence* Norwood, NJ: Ablex.

Campione, J. C., Brown, A. L., & Ferrara, R. A. (1982). Mental retardation and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence.* Cambridge: Cambridge University Press.

Campione, J.C., Brown, A.L., Ferrara, R.A., Jones, R.S., & Sternberg, E. (1985). Breakdown in flexible use of information: Intelligence-related differences in transfer folowing equivalent learning performance. *Intelligence, 9,* 297-315

Camplloch, J.M. (1981). Evaluación del retraso mental [Psychological assessment of Mental Retardation]. In R. Fernández-Ballesteros & J. A. Carrobles (Eds.), *Evaluación conductual* [Behavioral assessment]. Madrid: Pirámide.

Campllonch, J.M., & Fernández-Ballesteros, R. (1981). La evaluación del potencial de aprendizaje. [Learning potential assessment]. In R. Fernández-Ballesteros (Ed.), *Nuevas aportaciones en evaluación conductual* [Progress in behavioral assessment]. Valencia: Alfaplus.

Carlson, J.S., & Wieldl, K.H. (1978). The use of testing-the-limits procedures in the assessment of intellectual capabilities in children with learning difficulties. *American Journal of Mental Deficiency, 82,* 559-564.

Carpenter, P.A., Just, M.A., & Shell, D. (1990). What intelligence test measure: A theoretical account of the processing in the Raven Progressive Matrixes Test. *Psychological Review, 97,* 404-431

Carroll, J. B. (1993). *Human cognitive abilities. A survey of factor-analytic studies.* Cambridge: Cambridge University Press.

Catania, A.C., Matheus, B.A., & Shimoff, E.H. (1990). Properties of rule governed behavior and their implications. In D.E. Blackman & H. Levine (Eds.), *Behavior analysis in theory and practice.* London: LEA.

Cattell, R. B. (1971). *Abilities: Their structure, growth and action.* Boston: Houghton-Miflin.

Cattell, R.B. (1975). *Test de factor G* [Factor G test]. Madrid: TEA.

Cattell, R. B. (1987). *Intelligence: Its structure, growth and action.* Amsterdam: North-Holland.

Colom Maranón, B. R. (1989). Las Estrategias diferenciales de representación en el razonamiento [Differential strategies of representation in reasoning]. Unpublished dissertation. Universidad Autónoma de Madrid.

Colom Maranón, B. R., & Juan-Espinosa, M. (1990). *Estudios sobre los fundamentos de la Ccgnición* [Studies on the foundations of cognition]. Valencia: Promolibro.

Colom Maranón, B. R., Alonso, S., & Blanco, M.A. (1991). *Diferencias individuales en la resolución de inferencias transitivas: Demandas de procesamiento verbal y viso-espacial* [Individual differences when solving trasitive inferences: Verbal and spatial processing requierements]. Technical report. Universidad Autónoma de Madrid.

Colom Maranón, B. R., Blanco, M.A., & Alonso, S. (1990). *Diferencias individuales en la resolución de inferencias transitivas: Estrategias a partir de un análisis de protocolos verbales* [Individual differences when solving trasitive inferences: Strategies behind verbal reports]. Technical Report. Universidad Autónoma de Madrid.

Colom Maranón, B.R. (1995). *Psicología de las Diferencias Individuales* [Psychology of Individual Differences] (second edition). Madrid: Pirámide.

Colom Maranón, B.R., & Alonso, S. (1993). *Estrategias de la memoria de trabajo y asimetria hemisférico-funcional* [Working memory strategies and brain assymetry]. Technical Report. Universidad Autónoma de Madrid.

Colom Maranón, B.R., Blanco, F., & Ferreirós, J. (1992). *Bosquejo para un nuevo programa para el análisis del sistema cognitivo* [Sketch for a new program for the analysis of the cognitive system]. IV Seminario Interuniversitario de Ciencia Cognitiva. Valencia, 16-18 de Enero.

Cooper, L. (1982). Strategies for visual comparison and representation: Individual differences. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1). Hillsdale, NJ: LEA.

Cooper, L., & Mumaw, R. (1985). Spatial aptitude. In R. Dillon (Ed.), *Individual differences in cognition* (Vol.2). London: Academic Press.

Cornelius, S.W., & Caspi, A. (1987). Everyday problem solving in adulthood and old age. *Psychology and Aging, 8,* 144-151

De Bono (1976). *Teaching thinking*. London: Peguin .

Díaz Fernández, O. (1982). *Evaluación del potencial de aprendizaje* [Learning potential assessment]. Unpublished dissertation. Universidad de Santiago.

Downs, R.M. (1981). Maps and mapping as metaphors for spatial representation. In L.S. Liben (Eds.), *Spatial representation and behavior across life span: Theory and application*. New York: Academic Press.

Downs, R.M., & Stea, D. (1973). Cognitive maps and spatial and spatial behavior. Process and products. In R.M. Downs & D. Stea (Eds.), *Image and environment*. Chicago: Aldine Publishing Company.

Eliot, J., & Smith, I.M. (1983). *An international directory of spatial tests*. Windsor, England: Humanities Press.

Ettiene, A.S. (1992). Navigation of a small mammal by dead reckoning and local cues. *Current Directions in Psychological Science, 1,* 48-52.

Eysenck, H. J., & Barret, P. (1985). Psychophysiology and the measurement of intelligence. In C. R. Reynolds, & V. L. Willson (Eds.), *Methodological and statistical advances in the study of individual differences*. New York: Plenum Press.

Eysenck, H.J. (1979). *Structure and measurement of intelligence*. Berlin: Springer-Verlag.

Ferguson, G.A. (1956). On transfer and the abilities of man. *Canadian Journal of Psychology, 10,* 121-131.

Fernández-Ballesteros, R. (1979). Nuevas perspectivas en Psicodiagnóstico. Un ejemplo: la evaluación del potencial de aprendizaje [Advances in psychodiagnostic. An example: learning potencial assessment]. In A. Dosil (Ed.), *Desarrollo humano* [Human development]. Servicio de Publicaciones, Universidad de Santiago.

Fernández-Ballesteros, R. (1982). *Evaluación de Contextos* [Context assessment]. Murcia: Servicio de Publicaciones, Universidad de Murcia.

Fernández-Ballesteros, R., & Calero, M.D. (1992). Nuevos enfoques en la medida de la intenigencia: La evaluación del potencial de aprendizaje [New approaches in the measurement of intelligence: Learning potential assessment]. In *Inteligencia y cognición* (pp. 546-590) [Intelligence and cognition]. Madrid: Complutense.

Fernández-Ballesteros, R., & Calero, M.D. (1993). Measuring learning potential. *International Journal of Cognitive Education & Mediated Learning, 3,* 1, 9-21.

Fernández-Ballesteros, R., & Calero, M.D. (1995). Training effects on intelligence of older persons. *Archives of Gerontology and Geriatrics, 20,* 135-148.

Fernández-Ballesteros, R., Calero, M.D., Campllonch, J.M., & Belchi (1987). *Evaluación del potencial de aprendizaje* [Learning potential assessment]. Madrid: MEPSA.

Fernández-Ballesteros, R., Calero, M.D., Campllonch, J.M., & Belchi (1990). *Evaluación del potencial de aprendizaje* [Learning potential assessment]. Madrid: MEPSA.

Fernández-Ballesteros, R., Camplloch, J.M., & Maciá, A. (1982). Generalización en el entrenamiento de estretegias cognitivas [Generalization in the training of cognitive strategies]. In R. Fernández Ballesteros (Ed.), *Evaluación de contextos* [Context assessment]. Murcia: Sevicio de Publicaciones de la Universidad de Murcia.

Ferrara, R.A., Brown, A.L., & Campione, J.C. (1986). Children's learning and transfer of inductive reasoning rules: Studies in proximal development. *Child Development, 57,* 1987-1099.

Feuerstein, R. (1979). Ontogeny of learning man. In M.A.B. Braizier (Ed.), *Brain mechanisms in memory and learning from the single neuron to man.* New York: Raven Press.

Feuerstein, R. (1979). *The dynamic assessment of retarded performers: the learning potential assessment device, theory, instruments, and techniques.* Baltimore: University Park Press.

Feuerstein, R., Rand, Y., Hoffman, M., & Miller, R. (1980). *Instrumental enrichment. An intervention program for cognitive modifiability.* Baltimore, University Park Press.

Feuerstein, R., Haywood, C., Rand, Y., Hoffman, M., & Kensen, M.R. (1989). *Learning Potential Assessment Device Manual* Jerusalem: Hadassah-Wizo-Canada Institute

Feuerstein, R., Krasilowsky, D., & Rand, Y. (1974). Innovative educational strategies for the integration of high-risk adolescents in Isarael. *Phi Delta, Kappan, 8,* 1-6.

Feuerstein, R.Y., Rand, M., Hoffman, M., Hoffmand, M., & Miller, R. (1979). Cognitive modifiability in retarded adolescents: Effects of instrumental enrichment. *American Journal of Mental Deficiency, 83,* 539-550.

García, O. (1993). *Diferencias Individuales en Tareas de Memoria y Demandas de Procesamiento: un estudio de patrones de ejecución* [Individual differences in memory and processing requierements: A study of performance patterns]. Unpublished dissertation. Universidad Aut"noma de Madrid.

García, O., & Colom Maranón, B. R. (In press). Memorizadores lingüísticos y memorizadores visuales: Diferencias individuales en tareas de memoria operativa [Linguistic and visual memorizers: Individual differences in working memory tasks]. *Revista de Psicología General y Aplicada.*

Gould (1981). *The mismeasure of man.* New York: Norton.

Guilford, J.P., & Lacy, (1947). *Printed classification tests.* Army Air Force Aviation Psychology Program Research Reports, No. 5. Washington, DC: U.S. Government Pinting Office.

Guilford, J.P., & Zimmerman, E.S. (1948). The Guilford-Zimmerman aptitude survey. *Journal of Applied Psychology, 32* (1).

Gustafsson, J. (1984). A unifying model for the structure of intellectual abilities. *Intelligence, 8,* 179-203.

Guthke (1982). The learning test concept—An alternative to the traditional static intelligence test. *The German Journal of Psychology, 6,* 306-324.

Guttman, L. (1965). *The structure of interrelations among intelligence tests.* Actas del Congreso de 1964 sobre problemas de los tests. Princeton, NJ: Educational Testing Service.

Haywood, H.C, & Izuriel, D. (1992). *Interactive assessment* New York: Springer-Verlag.

Haywood, H.C. (1970). *Socio-cultural aspects of mental retardation.* New York: Appleton-Century-Crofts.

Haywood, H.C., Filler, J.W., Shiffman, M.A., & Chatelant, G. (1975). Behavioral assessment in mental retardation. In P. McReynolds (Ed.), *Advances in psychological assessment* (Vol 3). San Francisco, CA: Jossey-Bass.

Hoffman, K., Guilford, J.P., Hoepfner, R., & Doherty, W. (1968). A factor analysis of the figural-cognition and figural-evaluation abilities. Los Angeles: Reports from the Psychological Laboratory, University of Southern California, No. 40.

Hunt, E. (1974). Quote the Raven? Never more! In L.W. Greegg (Ed.), *Knowledge and cognition* (pp. 129-157). New York: Wiley and Sons.

Hunt, E., (1978). Mechanics of verbal ability. In P. Vernon (Ed.), *Speed of information-processing and intelligence.* Norwood, NJ: Ablex.

Hunt, E. (1980). Intelligence as an information processing concept. *British Journal of Psychology, 71,* 449-474.

Icarán, E., Colom Maranón, B.R., & Martín-Loeches, M. (1993). *Diferencias individuales en las estrategias cognitivas utilizadas ante la tarea de verificación de frases, y cartografía cerebral*

[Individual differences in cognitive strategies used solving sentence verification tasks and brain imaging]. Technical report. Universidad Autónoma de Madrid.

Infield, S.E., Yee, P.L., & Hunt, E.B. (1990). *A psychometric piecture of navigational abilities.* Paper presented at the meeting of the Ameriacan Psychological Society. Dallas, Texas.

Juan-Espinosa, M. (1994a). Urban and rural implicit theories of intelligence. *Proceedings of the XII Congress of Cross-cultural Psychology*. Pamplona.

Juan-Espinosa, M. (1994b). People's conceptions of apatial orientation ability: The case of "fang-okak" hunters of Equatorial Guinea. *Proceedings of the XII Congress of Cross-cultural Psychology*. Pamplona.

Juan-Espinosa, M. (In press a). *Psicología diferencial de la inteligencia* [Differential psychology of intelligence]. Madrid: Pirámide.

Juan-Espinosa, M. (In press b). *Psicología cognitiva diferencial* [Differential cognitive psychology]. Madrid: Pirámide.

Juan-Espinosa, M., & Colom, R. (1989a). El marco conceptual de la psicología cognitiva diferencial [The conceptual framework of the differential cognitive psychology]. In M. Juan-Espinosa & R. Colom (Eds.), *Psicología diferencial y cognición* [Differential psychology and cognition]. Valencia: Promolibro.

Juan-Espinosa, M., & Colom, R. (1989a). El marco de investigación de la psicologRaia cognitiva diferencial [The framework of research of the Differential Cognitive Psychology]. In M. Juan-Espinosa & R. Colom (Eds.), *Psicología diferencial y cognición* [Differential psychology and cognition]). Valencia: Promolibro.

Just, M.A., & Carpenter, P.A. (1976). Eye fixations and cognitive processes. *Cognitive psychology, 8*, 441-480.

Karpov, Y.V., & Talyzina, N.F. (1986). Criterios para el diagnóstico: Desarrollo intelectual [Diagnostic criteria in the development of intelligence]. *Evaluaci"n psicológica/ Psychological Assessment, 2,* 4, 3-17.

Keating, D.P. (1978). A search of social intelligence. *Journal of Educational Psychology, 70,* 218-223.

Kliegl, R., Smith, J., & Baltes, P. (1989). Testing the limit and the study of adult age differences in cognitive plasticity of a nemonic skill. *Developmental Psychology*

Kosslyn, S., Ball, T.M., & Reiser, B.J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 1, 47-60.

Kyllonen, P. (In press). Entrevista realizada por B. R. Colom Maranón [Interview to Patrick C. Kyllonen]. *Estudios de Psicología.*

Kyllonen, P., & Christal, R. (1989). Cognitive modeling of learning abilities: A status report of LAMP. In R. Dillon & J. W. Pellegrino (Eds.), *Testing: theoretical and applied issues.* New York: Freeman.

Kyllonen, P., & Christal, R. (1990). Reasoning ability is (little more than) working-memory capacity? *Intelligence, 14*, 389-433.

Kyllonen, P., Lohman, D., & Snow, R. (1984a). Effects of aptitudes, strategy training, and task facets on spatial task performance. *Journal of Educational Psychology, 76*, 130-145.

Kyllonen, P., Lohman, D., & Woltz, D. (1984b). Componential modeling of alternative strategies form performing spatial tasks. *Journal of Educational Psychology, 76*, 1325-1345.

Labouvie-vief, G. (1982). Dynamic development and mature autonomy: A theoretical Prologue. *Human Development, 25*, 161-191.

Leo, E.M. (1993). *Diferencias individuales en la tarea de inferencia transitiva: Estudio de patrones de ejecución con un paradigma de libre elección* [Individual differences when solving trasitive inferences: A study of performance patterns with a free-selection paradigm]. Unpublished dissertation. Universidad Autónoma de Madrid.

Leo, E.M., & Colom Maranón, B. R. (In press). Razonamiento lógico o psicológico? Diferencias individuales en una tarea de razonamiento deductivo usando un paradigma de libre elección [Logic or psychological reasoning? Individual differences when solving deductive reasoning tasks with a free-selection paradigm]. *Revista de Psicología General y Aplicada.*

Lidz, C.S. (1987a). Cognitive deficiencies revisited. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 444-478). New York: Guildford Press.

Lidz, C.S. (1987b). Historial perspective. In C.S. Lidz (Ed.), *Dynamic Assessment.* New York: Academic Press.

Lohman, D. (1988). Spatial abilities as traits, processes, and knowledge. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 4). Hillsdale, NJ: LEA.

Lohman, D.F. (1988). Spatial abilities as trait, processes, and knowledge. In R.J. Sternberg (Ed), *Advances in the psychology of human intelligence* (Vol. 4). Hillsdale, NJ: Lawrence Erlbaum.

Loomis, J.M., Klatzky, R.L., Golledge, R.G., Cicinelli, J.G., Pellegrino, J.W., & Fry, P.A. (1993). Nonvisual navigation by blind and sighted: Assessment of path integration ability. *Journal of Experimental Psychology: General, 122*,1, 73-91.

López, R. (1993). *Estudio de los patrones de ejecución de ciegos totales y videntes en tareas de inferencia transitiva* [A study of performance patterns of blind and sighted subjects when solving transitive inferences]. Unpublished dissertation. Universidad Autónoma de Madrid.

López, R., & Colom Maranón, B. R. (1995). Ciegos y videntes: influencia de las limitaciones estructurales al razonar [Blind and sighted subjects: the influence of structural limitation in reasoning]. *Integración, 17*, 5-11.

López, R., Leo, E.M., Blanco, F., & Colom Maranón, B. R. (1994). Exploración háptica y estimación mental de distancias en una isla ficticia: la importancia de la acción en las representaciones espaciales [Haptic exploration and mental exploration of distances within a fictious island: The importance of action to spatial representation]. *Integración, 14*, 5-15.

Luciano, M.C. (1992). La conducta verbal a la luz de recientes investigaciones. Su papel sobre otras conductas verbales y no verbales. *Psicothema, 4,* 445-468

MacCleod, C.M. (1991). *Half a century of research on the stroop effect: An integrative review.* Division of life sciences. University of Toronto, Scarborough campus, Scarborough, Ontario, Canada.

MacCleod, C.M., Hunt, E., & Mathews, N. (1978). Individual differences in the verification of sentence picture relationships. *Journal of Verbal Learning and Verbal Behavior, 5*, 493-508.

Mathews, N., Hunt, E., & MacCleod, C.M. (1980). Strategy choice and strategy training in sentence picture verification. *Journal of Verbal Learning and Verbal Behavior, 19*, 531-548.

Palacios, A. (1993). *Diferencias individuales en el razonamiento analógico. Un estudio basado en patrones de ejecución* [Individual differences in analogical reasoning: A study of performance patterns]. Unpublished dissertation. Universidad Autónoma de Madrid.

Palacios, A., & Colom Maranón, B. R. (In press). Razonadores lingüísticos y razonadores visuales: Diferencias individuales en el razonamiento analógico con un paradigma de libre elección [Visual and linguistic reasoners: Individual differecnes in analogical reasoning with a free seleciont paradigm]. *Revista de Psicología General y Aplicada.*

Pellegrino, J.W., & Hunt, E.B. (1991). Cognitive models for understanding and assessing spatial abilities. In H.A.H. Rowe (Ed.), *Intelligence: Reconceptualization and measurement.* Hillsdale, NJ. Lawrence Erlbaum Associates.

Raven, J.C. (1968). *Test de matrices progresivas. Escala especial* [Progressive matrices test. Special scale]. Madrid: MEPSA.

Raven. J.C. (1975). *Test de matrices progresivas. Escala general* [Progressive matrices test. General scale]. Madrid: MEPSA.

Risberg, J., & León-Carrión, J. (1988). La medida del flujo sanguineo cerebral regional en evaluaci"n neuropsicológica [Measurement of regional cerebral blood flow in neuropsychological assessment]. *Evaluación Psicológica/ Psychological Assessment, 4,* 339-352

Schaie, K.W. (1990). Intellectual development in adulthood. In J. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging*, Third Edition. New York: Academic Press

Siegel, A.W., & White, S.H. (1975). The development of spatial representation of large-space environments. In H.W. Reese (Ed.), *Advances in child development and behavior.* New York: Academic Press.

Snow, R., & Lohman, D. (1993): Cognitive psychology, new test design, and new test theory: An introduction. In N. Frederiksen, R.J. Mislevy, & I.I. Bejar (Eds.), *Test theory for a new generation of tests.* New Jersey: LEA.

Snow, R., & Swansson, J. (1992). Instructional psychology: Aptitude, adaptation, and assessment. *Annual Review of Psychology, 43*, 583-626.

Snow, R., Kyllonen, P., & Marshalek, B. (1984). The topography of ability and learning correlations. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2). Hillsdale, NJ: LEA.

Snow, R.E. (1989). Aptitude-treatment interaction as a framework for research on individual differences in learning. In P.L. Ackerman, R.J. Sternberg, & R. Glaser (Eds.), *Learning and individual differences: Advances in theory and research.* New York: Freeman.

Staats, A. W. (1968). *Learning, language, and cognition.* New York: Holt, Rinehart and Winston.

Staats, A.W. (1971). *Child, learning, intelligence, and personality*. New York: Harper & Row.

Staats, A.W. (1990). Paradigmatic behaviorism and intelligence: Task analysis?, technical plan? or theory? *Psicothema, 2,* 7-25

Sternberg, R.J. (1981). Testing and cognitive psychology. *American Psychologists, 36,* 1181-1189

Sternberg, R. J. (1980). Sketch of a componential subtheory of human intelligence. *Behavioral and Brain Sciences, 3,* 573-614.

Sternberg, R. J. (1990). *Metaphors of mind: Conceptions of the nature of intelligence.* Cambridge: University Press.

Sternberg, R. J., Conway, B. E., Ketron, J. L., & Bernstein, M. (1982). People's conceptions of Intelligence. *Journal of Personality and Social Psychology, 41*, 37-55.

Sternberg, R.J. (1977). Components processes in analogical reasoning. *Psychological Review, 84,* 353-378.

Sternberg, R.J. (1979). A review of "Six authors in search of a charater": A play about intelligence tests in the year 2000. In R.J. Sternberg & D.K. Detterman (Eds.), *Human intelligence.* Norwood, NJ: Ablex.

TEA. (1971). *Test de aptitudes mentales primarias* [Primary mental abilities -PMA- test]. Madrid: Tea Ediciones.

TEA. (1976). *Test de inteligencia general. Serie Dominós, forma I* [General intelligence test: domino series, Form I]. Madrid: Tea Ediciones.

Thorndike, E.L. (1920). Intelligence and its uses. *Harper's Magazine, 140,* 227-235.

Vernon, P. E. (1979). *Intelligence, heredity and environment.* New York: W.H. Freeman.

Vernon, P. E. (1969). *Intelligence and cultural environment.* London: Mathuen.

Vye, N.J., Burns, M.S, Declós, V.R., & Bransford (1987). Dynamic assessment of intellectually handicapped children. In C.S. Lidz (Ed.), *Dynamic assessment.* New York: Guildford Press.

Vygostky, L.S. (1934/ 1962). *Thought and language.* Cambridge, MA: MIT Press.

Vygostky, L.S. (1935/ 1978) *El desarrollo de los procesos psicológicos superiores* [The development of higher psychological processes]. Barcelona: Grijalbo.

Wechsler, D. (1974). *Escala de inteligencia de Wechsler para ninos. Manual* [Wechsler Intelligence Scale for Children. Manual]. Madrid: TEA Ediciones.

Willis, S.L., Blieszner, R., & Baltes P.B. (1981). Intellectual training research in aging: Modification of performance on thge fluid ability of figural relations. *Journal of Educational Psychology, 73,* 41-50.

Willis, S.L. (1987). Cognitive training and everyday competence. *Annual Review of Gerontology and Geriatrics, 7,* 159-188.

Witkin, H. A. (1967). A cognitive-style approach to cross-cultural research. *International Journal of Psychology.*

Yuste. M. (1983). *Fichas de entrenamiento en orientación espacial* [Spatial orientation training materials]. Madrid: CEPE.

ON THE BOUNDARIES OF THE ACCELERATION OF THE DEVELOPMENT OF INTELLIGENCE

Johannes Kingma and Welko Tomic

INTRODUCTION

The development of intelligence refers to the changes that an individual's mental apparatus undergoes due to experiences during his or her life cycle. According to Piaget, the development of intelligence takes place in a sequence of discrete stages. A child's way of thinking in one particular stage is qualitatively different than in the previous or subsequent phases.

With regard to intelligence, we can distinguish between cognitive structure and function. Cognitive structure is a hypothetical mental construct that changes throughout development. Function concerns the internal and external actions related to the structure.

The development of intelligence refers to a genesis of structures. Intelligence develops by refining and transforming mental structures (Piaget, 1963, 1964a, 1964b, 1973). Structure as defined by Piaget means an organized totality within

Advances in Cognition and Educational Practice, Volume 4, pages 275-290.

ISBN: 0-7623-0105-8

which the relationships between elements are clearly defined. Cognitive structures refer to mental knowledge and production systems that are not directly observable but that lie at the basis of intelligent actions. In simplified form, a structure can be seen as a type of knowledge database that a child uses to interpret the world. The child knows the world or observes reality in terms of its structures. Piaget attempts to describe and explain the development of intelligence by postulating general abstract structures in which seemingly different intelligent actions might be arranged. Successive structures therefore differ qualitatively throughout the course of the development of intelligence.

Function concerns the internal and external actions related to the structure. Activating a cognitive process is an example of an internal action. For example, in attempting to recall something, we are activating a series of internal actions that retrieve that particular word or concept from memory. Experience is the external aspect of function: it is the external source of stimulation.

Development is often viewed as the interaction between structure and function. Activities that take place both in the environment and in the structure itself can contribute to changes in the structure, which in turn bring about changes in the way the structure operates.

Piaget does not view the development of intelligence as a gradual accumulation of knowledge or skills, but rather as a sequence of structural transformations: rather abrupt, intermittent changes in the way a child thinks. During the transition between two stages, the child inhabits two qualitatively different worlds. The discrepancy between what things seem to be and what they really are leads to a conflict in the child's way of thinking. His thinking has progressed to a new and higher level and is now qualitatively different from his thinking based on the previous structure.

Another aspect of Piaget's theory is the notion that the activity of children is intrinsic. Their structures are intrinsically active, intrinsically curious. Children are not satisfied with what they already know, but are constantly in search of greater knowledge. The motivation to develop is generated from within. Although Piaget acknowledges that environmental and biological factors play a role, he considers intrinsic activity the motor of cognitive development. The fuel for this motor is the reciprocal relationship between function and structure, namely that the child's activity, or in reality the activity of the structures, influences the subsequent development of these structures. Intelligence develops as a result of the process of construction performed by the child. The child constructs his own reality. The child naturally interprets the information that reaches him from his surroundings in terms of the information he already possesses. Development of intelligence means a change in both knowledge and ability. The database changes and consequently so do its structure and organization, and this alteration produces a change in how the child perceives reality.

According to Piaget, education must keep pace with the actual intelligence level of a child, so that the child can understand (assimilate) the new information. It is the conflict between the appearance of an object and what it actually is that in fact induces a change in the structure. The child's intrinsic activity and motivation together form the dynamic motor of the process that will lead to the resolution of such a conflict. The most suitable approach to learning is the method of self-discovery, in which the teacher remains in the background, offers suitable material at the appropriate time, and continuously asks questions that will encourage the child to justify his solutions. The teacher does not provide any feedback whatsoever.

Piaget is rather skeptical about the possibility of accelerating the development of intelligence through training or other interventions. Only children who already possess partial knowledge of the concept in which they are to be trained can benefit from training.

This chapter will review an important intelligence issue related to the Piagetian view. We will explore whether and if so, to what extent, the development of intelligence can be accelerated. This question is important in both theoretical and practical terms. The chapter concludes by discussing the different research results, and puts them into a theoretical perspective.

CAN THE DEVELOPMENT OF INTELLIGENCE BE ACCELERATED?

This section begins by describing the Genevan and American training studies. The discussion then turns to the various training methodologies and describes the attempts made to accelerate the development of intelligence. Finally, the authors question why it is considered desirable to accelerate the development of intelligence.

Genevan Training Research

If adaptation is assumed to be a basic function of intelligence and of a child's self-regulating activity, one can understand Piaget's position on the possibility of accelerating the development of intelligence. The interpretation of Piaget's position can be described briefly in the following terms: "Readiness to learn." Not only should new subjects match the cognitive level at which a child presently functions, but learning itself should be embedded in the child's self-regulating activity. Wadsworth (1978) aptly describes this twofold condition in the following way:

> The function of the teacher is not to accelerate the development of the child or speed up the rate of movement from stage to stage. The function of the teacher is to insure

> that development within each stage is thoroughly integrated and complete (Wadsworth, 1978, p. 117).

Socio-educational transfer or instruction in knowledge and skills only makes sense when it links up with the cognitive structure that already exists. If we try to teach a child something that it is not capable of learning yet, it will probably be able to solve the problem it has been trained to solve, but no more than that. It will be unable to generalize enough to solve similar but slightly different problems, even after a period of instruction. The phenomenon of transfer, that is, the ability to apply what one has learned in other training situations, does not take place (Kingma & TenVergert, 1993b, 1993c, 1993d).

For Piaget, it is precisely the range of the transfer which indicates whether instruction or training has brought about a fundamental change in a child's cognitive structure. If the child is unable to generalize, then training has in all likelihood generated an isolated structure, so that the child is unable to use the principle it has learned to solve other problems.

Nonetheless, Piaget (1959a, 1959b) believes that it is possible to accelerate the development of intelligence by using certain specific instruction methods. That is only possible when instruction results in more complex cognitive structures being generated from simpler ones. Piaget therefore has a number of stringent requirements for both the method of instruction and the results achieved. Regarding the instruction method, the training methodology must correspond to the spontaneous development of intelligence (Inhelder, Sinclair, & Bovet, 1974). According to Piaget and his colleagues, the interview is the most suitable technique and, hence, the most important method of instruction, for example, for stimulating the construction principle (i.e., the manipulation of objects) in children.

> By observing and interviewing children..., teachers gain empathy and respect for children's developing intellectual capacities. Such an understanding of their intellectual processes prevents the teacher from teaching children concepts they are not prepared to learn. In other words, this teacher is aware of the natural constraints of the child's stage of development. At the same time, the teacher has respect for the child's current capacity for learning and is aware of the multiplicity of new capacities that become available at each new level of development. This level of awareness alerts the teacher to curriculum materials that place artificial constraints on the children's natural capacities and provides her with a basis for making on-the-spot curriculum decisions in the classroom (Labinowicz, 1980, p. 160).

The methodology described in this quotation demonstrates that instruction of training derived from Piaget's theory is viewed as follows: during instruction, the child is given a range of objects to manipulate. When the child solves a problem the teacher or experimenter interviews the child, for example by asking: "Why is that so?" or "Do you think that's correct?" or "How do you know?" One of the fundamental principles of the Genevan method of instruction is that

the child must discover the method of solution himself by manipulating objects. The self-discovery method, according to Piaget, is the most suitable way to teach or train intelligent actions, because it is a method that complements the child's own self-regulating activity. That brings us back to the starting situation or the child's actual cognitive level, which determines how successful instruction can in fact be. After all, children who are in a period of transition will experience a conflict between what *is* (direct observation) and what *might be* (it could be true). They will intrinsically be inclined to resolve that conflict themselves. The "*verstehend*" (empathic) capacity of the teacher or experimenter is of primary importance when it comes to creating situations for the child and offering the child the opportunity to resolve the conflict on his own.

It only makes sense to teach a new concept if the child already possesses partial knowledge of that concept. Children who are entirely unfamiliar with the concept will not benefit from training or instruction (Piaget, 1964a, 1964b). A child must first have the cognitive structure into which the new information can be assimilated.

> This is why we cannot teach higher mathematics to a five-year-old. He does not yet have structures which enable him to understand" (Piaget, 1964b, p. 13).

Piaget's colleagues undertook numerous training experiments in the seventies in which they attempted to show that successful training depends on the child already having some understanding of a concept or of intelligent solution methods (Inhelder et al., 1974). In Piaget's opinion, a successful "learning achievement" has to satisfy three requirements. First, the training or learning effect has to be evaluated from the perspective of the spontaneous development of intelligence. Second, there has to be evidence of transfer to conceptual areas in which the child has not been trained. For example, if the child has been trained to solve problems involving conservation of length, after training it should also be able to solve other, related problems, for example, involving seriation and transitivity of length. Finally, the learning effect has to be durable. The child should be re-tested on the problems listed under item 2 a few weeks to a few months (as many as four) later. If he is still able to solve these problems correctly even after a period of that length has passed, then training can be assumed to have been successful. Only if the three requirements are fulfilled, can one conclude whether training has induced a change in the child's cognitive structure.

The central issue of the second requirement is to convert the theoretical requirement into measurable terms. The general methodology of the Genevan school runs as follows. Prior to training, the children are administered a pretest in which they are asked to solve problems covering the subjects in which they are to be trained, for example conservation of length. Using Piaget's criteria, the children are, for instance, divided into non-conservers and partial conservers

based on their solutions and arguments. The experimenter, also the instructor, then begins the training program. Usually, the instructor trains only one child at a time. The Genevan training approach is not standardized, as it takes place by means of the interview method described previously. In standardized training methods, all the children are given the same treatment according to a previously determined procedure. The Genevan training course generally lasts two to three sessions of fifteen to twenty minutes each. After training, each child is administered a posttest which consists of problems that allow the researcher to assess the success of training, in accordance with Piaget's second requirement. The children's solutions to the posttest problems also indicate the range of transfer, which determines whether or not training has been successful. The child's score on the first posttest, administered immediately after training, is not the only determining factor, however. The "actual" level of intelligent action can be estimated. In a second, and even a third or fourth posttest, the children are asked to solve problems from the first posttest again, so that the durability of the learning and transfer effect can be determined. The results of Piaget's training experiments clearly show that children whose pretest indicates that they have no understanding whatsoever of the concept scarcely benefit from training, if at all (Inhelder et al., 1974). Even if the results of the immediate posttest show that the children have made some progress over the pretest results, there is still a very good chance that they will drop back to their old level by the time they take the second or third posttest.

The crucial question is whether it is possible to accelerate the development of intelligence in a durable way and at a higher level by means of instruction or training. The answer is yes, although with a few serious reservations. From the Piagetian perspective, the cognitive structure must be shaped in such a way that it can assimilate the new information. The possibility of acceleration is therefore restricted to those children who already have a partial knowledge or some notion of the concept in which they are to be trained. The training methodology must be based on the child's own self-regulatory function. In addition, there are *stringent* requirements with respect to the training effect. Piaget's training results show that short-term training will only accelerate the development of intelligence to a very limited extent. With this in mind, we may question whether it is indeed generally possible to raise children to a higher level of intelligence through training. Piaget's position on this is rather pessimistic: short-term training is unsuitable. The following quote by Woolfolk (1987) is relevant here.

> If you try to teach a student something the student is not ready to learn, he or she may learn to give the 'correct' answer. But this will not really affect the way the student thinks about this problem or any other problem. Therefore, from this perspective acceleration is useless. A second Piagetian argument is that acceleration is inefficient. Why spend a long time teaching something at one stage when the students will learn it by themselves much more rapidly and thoroughly at another stage (Woolfolk, 1987, p. 70)

American Training Research

In the early 1960s, a major dispute erupted specifically between the Genevan school and American researchers. The controversy was sparked by Bruner's (1960) *The Process of Education*, which focused on the issue of "Readiness to learn." Bruner disagreed with Piaget's idea that education had to keep pace with the child's actual cognitive level. In his view, it was precisely the task of education to anticipate that structure in order to stimulate cognitive development. Bruner influenced numerous American researchers, leading to a veritable boom in training experiments. These experiments focused largely on children who had not yet acquired any notion of the concept in which they were to be trained or of the solution method, for example non-conservers with respect to number, quantity, and so on. It is only a slight exaggeration to say that the researchers conducted extremely simple training experiments. While working on a conservation problem, for example, the experimenter would tell a non-conserver child "You can pour the water back, can't you?" or even demonstrate it. The posttest would then be administered, with the experimenter using orange juice instead of water, for example. Some training sessions focused on only one problem, and the sessions sometimes varied from two to twenty or even forty minutes (Kingma, 1981). If we apply the three Piagetian criteria to the learning effects of the American training research, it becomes clear that these effects are relatively specific and short term, with only very limited transfer taking place. There was no research on durability. According to Kingma and Koops (1988), the American researchers' behaviorist background lay at the root of their unlimited optimism with respect to accelerating the development of intelligence in children who had not yet acquired any notion of the concept in which they were to be trained. Despite the very weak empirical evidence, the results apparently satisfied those researchers, but their findings do not support Piaget's theory on accelerating the development of intelligence by means of short-term training (Tomic, 1995a).

By now it has in fact been demonstrated that long-term training (i.e., lasting three to four weeks, during which, for instance, non-conservers and non-seriators receive fifteen to twenty minutes of training every school day) can generate massive training effects that meet Piaget's criteria and are observable for more than four months (Kingma, 1984b, 1986; Kingma & Loth, 1983, 1984; Kingma & Koops, 1984a, 1984b; Tomic, Kingma, & TenVergert, 1993). Long-term training can have a decisively positive impact on accelerating the development of intelligence.

Various Training Methods

It is important to note that various training methods are suitable to accelerate the development of intelligence. For example, Kingma (1986) constructed a

training program on seriation based on the American preschool curriculum developed by Hooper and Marshall (1968) and Hooper (1973). Tomic et al. (1993) designed a training program on measurement based on a Soviet method developed by Obuchova (1966, 1972). The program applies the theoretical model of the stepwise formation of mental action (Gal'perin, 1966) in a practical classroom situation. Tomic and Klauer (1996) have used a paradigmatic training method in studies focusing on inductive reasoning and problem-solving. These three programs, each of which was based on a different theory, accelerated the development of intelligence and met Piaget's stringent criteria for determining the success of training.

After the "success" of the American training studies in the sixties, Beilin (1971) came to the conclusion that every type of training program is effective when it comes to learning the concepts underlying Piaget's problems. His conclusion is drawn from other, and weaker, assessment criteria than those formulated by Piaget. When Piaget's stringent criteria are applied, more than 95 percent of the training studies in fact did not succeed in accelerating the development of intelligence (Kingma, 1981). The relatively small number of training programs whose results do meet Piaget's Stringent criteria are based on various theoretically founded theories of instruction. A theoretically derived methodology is evidently not the panacea which can accelerate the development of intelligence. On the contrary, more general aspects, present in every successful method, are decisive. In each successful method, the curriculum is constructed systematically. For example, suitable curriculum sequences are determined, the partial actions which underlie the training concepts are taught systematically, and regular repetition of these partial actions in various situations are spread out over a period of several weeks or more. The tasks or problems bring about a cognitive conflict in the child. The child has not acquired enough experience to solve the problem or to resolve the cognitive conflict. The child is hence guided in such a manner that a new, generally more complex cognitive structure arises. The scope of the transfer which has been generated is broad. Finally, an attempt is made to make the child aware of the new method of reasoning. This metacognitive aspect is very important. These are features of successful methods that can also be found in Adey and Shayer (1995).

A Permanently Higher Level of Intelligence?

Over time a wide range of interventions have been applied in the form of training courses. Beilin (1971) and Kingma (1981) reported on many different small-scale experiments. Both researchers focused on what Piaget called "the American question": accelerating the development of intelligence through specific short training programs (Neimark, 1975). Large-scale training programs have been described and evaluated by Nickerson, Perkins and Smith (1985), Coles and Robinson (1989) and Nisbet and McGuiness (1990), and

compensation and (school) enrichment programs of a similar scale, such as Head Start, have been described by Carter (1984), Nurss and Hodges (1982), and Scarr, Weinberg and Levine (1986).

The findings drawn specifically from small-scale experiments and experiments carried out within larger-scale training programs and, to a lesser extent, compensation and school enrichment programs, might lead to the optimistic conclusion that training and education are a relatively easy way of getting children to operate *permanently* at a higher level of intelligence. This optimism must be tempered somewhat, given the magnitude of the very long-term effects of both training and school enrichment programs. The key question is: after a successful course of training which also meets Piaget's criteria, what happens to the advantage that the trained children have over non-trained children if training is not continued at a higher level? After all, after the course of training ends, the child returns to its original environment in mainstream education along with the rest. After a while, for example a year, the non-trained children in the control group will progress "naturally" to the same level as the trained children. On the other hand, there is a chance that the trained children will drop down to the level of the non-trained children once again.

We can illustrate the latter by looking at a successful method for teaching non-conserver children to solve conservation problems. The children are trained in measuring various quantitative aspects of objects, such as length, volume and surface area (Kingma & Loth, 1984; Tomic et al., 1993). The training programs induced strong and broad transfer effects, which continued to be observed four months after the training program had ended. These programs therefore more than satisfied Piaget's stringent requirements with respect to training success. After the training program, the trained children returned to their old classrooms. They were no longer given special training in specific measurement skills at a higher level. Two years after completing the training program, the performance of the trained children was slightly below that of two years before. What was striking was that the non-trained children had achieved approximately the same level as their trained classmates after two years. It appears that even without specific training, children are capable of learning concepts and solution strategies which allow them to solve problems taken from the training program accurately.

The results of very long-term training programs, some of which last from kindergarten to the final grade of primary school, also show that trained children enjoy a distinct advantage. The flip side of this remarkable success is that the majority of these children drop back down to the level of their untrained classmates after the extra training efforts cease. That is frequently also the result of social and psychological factors. This summary is based on reports on enrichment programs involving children from underprivileged socio-economic backgrounds, which were conducted in the sixties and seventies.

One of the most famous of these was the Head Start program, which focused on the theme "learning to learn":

> The curriculum usually included instruction in standard English and the use of basic concepts. Children were taught to classify objects and events in terms of temporal and spatial relationships 'before' and 'after'; 'distant' and 'near'), size, color, and shape. They were taught to count. Some programs taught reading as well. Instruction was offered on an individual or small-group basis, and children were allowed to proceed at their own pace. Head Start centers were not like first-grade classrooms, but neither were they like home. Educators spent a good deal of effort instilling such habits as paying attention, listening to teachers, and following the rules (Scarr, Weinberg, & Levine, 1986, p. 245).

The parents of participating children were involved in activities at centers which ran the pre-school training programs. The level of participation varied from one center to the next..

> The centers used a number of different models of parental involvement. Some programs were center based. In these, parents were informed about the program and, in some cases, participated in the planning; but they were not involved in daily activities. Other programs were home based. Here, 'parent educators' visited home; provided books, games and other materials for use in the home; and trained parents (usually mothers) in techniques for promoting cognitive development (Scarr, Weinberg, & Levine, 1986, p. 245).

The children who participated in Head Start programs clearly did not always receive the same instruction and/or training. The volume and type of training, as well as the degree to which parents were involved, varied both within and between the various centers. The unsystematic design meant that it was impossible to find a clear-cut explanation for any positive effects on the development of intelligence.

After Head Start, which was launched in 1965, had been running for a few years, the results of various studies showed that this type of education had a negative impact on the development of intelligence (Caruso, Taylor, & Detterman, 1982). Cicirelli (1969) showed that children who participated in the Head Start program did better at school than the control group of classmates, who had not had "preschool training." On closer analysis, however, it became clear that the advantage in terms of IQ and grades was relatively minor and that even those gains disappeared after two or three years in primary school. Other studies confirmed these findings (Scarr et al., 1986).

> Compensatory programs appear to result in immediate gains which are greater for target children than for control children. These broad gains appear to dissipate over time, however, so that scores for control and experimental children approximate both one another and their pre-intervention scores after a few years in elementary school (Nurss & Hodges, 1982, p. 50).

Other research demonstrated that compensatory programs like Head Start generally brought about short-term positive effects in better pupils (Carter, 1984). The effects of these programs are also biggest in the lower grades of primary school. Towards the end of primary school, however, any advantage acquired through a compensatory program has disappeared. Indeed, the training effect of long-term compensatory programs does not, in the long run, give even successful, gifted children a permanent advantage in the development of intelligence compared with classmates who did not participate in such a program.

Compensatory programs designed in the seventies and eighties were frequently criticized for not having clear-cut psychological and didactic underpinnings. Some "pre-school" curricula from that period were based on theoretical foundations, such as the "Piagetian Preschool Educational Program" (PPER). This program, implemented in the state of Wisconsin, involved three- and five-year-olds in day nurseries and kindergarten. Piaget's self-discovery method was one of the most important points of departure for the curriculum. Children were encouraged to actively manipulate objects and interact openly with their playmates. They were given neither correct solutions nor correct answers. For 28 weeks each school year, the children participated in a training session every school day. After three years, the participating children were evaluated. Compared with children in the control group who had not been trained, the trained children had made enormous progress on a whole series of Piagetian problems. Nevertheless, the control group children demonstrated similar progress (Hooper & De Frain, 1980, p. 172).

Weikart, Epstein, Schweinhant, and Bond (1978) showed that three different curricula, one of which was Piagetian, had the same effect on the development of intelligence. Although the three curricula were based on different theoretical foundations, these results showed that they had all successfully influenced the development of intelligence. Apparently, it is enough to systematically and regularly teach the relevant building blocks that will allow children to solve Piagetian problems. Even in cases where the curricula were more theoretically grounded, it became clear that without permanent follow-up after the training period had ended, the gains, in terms of accelerating the development of intelligence, were only temporary. Within a year or two, the non-trained children who continued their mainstream education were at the same level as the children involved in the experimental training conditions.

It could be stated, then, that a "long-term" training program lasting about four weeks is still too short to give the trained children a permanent advantage over their non-trained classmates in terms of their intelligence level. It should be noted that the criteria in that case are even more stringent than Piaget's.

The previous arguments make clear that the development of intelligence can be accelerated using specific training programs based on various curricula. The training effects of structured and long-term instruction appear to meet Piaget's stringent criteria with respect to training success. The fact that these positive

training effects dissipate in the long run, a year or two after the training program ends, shows that a number of environmental factors must also undergo permanent change if the benefits are to last.

We might, for example, consider the results of studies on how the environment influences the development of intelligence in monozygote twins. Normally, genetically identical twins are raised together by their biological parents. There are cases, however, where each twin is adopted by a different set of parents. It might be said that the study of identical twins who were raised in different environments resembles an experiment in which the independent variable, the environment, has been manipulated (Van der Zanden, 1985). If we compare the intelligence test results of a set of identical twins who were raised separately to that of twins raised in the same environment, it becomes possible to determine the relative impact of the environment. Summaries of correlational studies on monozygote twins (Bouchard & McGue, 1981; Hunt, 1961; Vernon, 1979) show that the average correlation coefficient on an intelligence test is approximately .90 for identical twins raised in the same environment. For identical twins raised in different environments, on the other hand, the correlation coefficient is approximately .80.

> If members of identical twin pairs had similar environments, these correlations may be interpreted as evidence that measured intelligence is at least partially determined by heredity. Related to these studies is the observation that with decreasing genetic similarity, there is a corresponding decrease in similarity between intelligence scores (Le Francois, 1984, p. 79).

The difference in the strength of the correlation (.90 - .80 = .10) may be attributable to environmental influences, among other things. Environment does indeed have an impact on the development of intelligence, but the genetic component nevertheless predominates.

Within the context of Piaget's theory, the environmental effect on the development of intelligence can be described in terms of the processes which underlie adaptation, that is, assimilation and accommodation.

Why Accelerate the Development of Intelligence

We are justified in questioning whether we should attempt to accelerate the development of intelligence if it turns out that children will acquire these concepts and solution strategies anyway, as part of a "natural" process of development, even if they do so later. The answer to this question can be broken down into three parts (Schmitt, 1994). First, human beings are persistent in their curiosity to know more. The question as to whether, and if so, to what extent the development of intelligence can be accelerated is simply an expression of this curiosity. Second, the idea of carrying out pioneering work,

conquering unknown territory and surmounting barriers is irresistible. Third, the point is not only to acquire knowledge, but knowledge that can be applied. The application of this knowledge must also be generally useful. This is an urge whose roots go back to sixteenth-century Western European philosophy. Francis Bacon (1561-1626), the spiritual father of the credo "knowledge is power," described how the laws of nature, those which had been discovered and which had still to be discovered, could be manipulated by instruments and tools. Nature could be tamed and put to work in the service of man: technology. It is, hence, precisely the results of training experiment interventions that can help the educational psychologist and the educationalist to develop instructional technology further.

> The practical importance of modifying intelligence is more obvious but independent from theoretical importance. The degree to which intelligence is thought to be a biological characteristic in theoretical models has no bearing on the practical importance of its alterability. All interventions must be environmental. The development of a technology of education depends on exactly understanding to achieve gains in intelligence (Detterman, 1982, p. VII).

The key question is whether the development of intelligence, as described in Piaget's theory, can be accelerated. The answer to the question is rather complex, but positive. Short training programs do not lead to a change in the cognitive structure and therefore do not induce an acceleration in the development of intelligence. On the other hand, long-term training programs do influence the development of intelligence. In the long run, however, a year or two after the program has ended, the positive effects dissipate. Studies of identical twins make clear that more permanent changes in the environmental factors influence the development of intelligence to a certain extent.

CONCLUDING REMARKS AND A LOOK AHEAD

Piaget's theory on the development of intelligence is related to the mental adaptation to new situations. A type of evolutionary process is repeated as the newborn develops into a young adult. After acquiring the formal operations, the person in question has all the cognitive tools he needs to be able to adapt to strongly fluctuating situations. Piaget bases his theory of stages, first of all, on observations of the behavior of very young children and second on the way in which children solve certain problems. The Piagetian tasks have an inherently theoretical reference point. Conservation tasks, for example, can help to determine whether a child is capable of reversable thinking within the conceptual area of the intended conservation problem. The range within which the child can perform reversible operations is then determined by means of conservation problems which cover different conceptual areas such as quantity, weight, circumference, distance, area, volume, and speed.

The heated debate on the possibility of accelerating intelligence appears to have burned itself out. The enormous number of training experiments frequently resulted in the trained children making statistically significant progress compared with the control group. Most of this progress can be attributed to the low standard used to evaluate training success. However, when Piaget's more stringent standards are applied, the majority of training effects were not successes at all. The conclusion is that the development of intelligence can in fact be influenced, regardless of the theoretical underpinnings. But lasting change is only effected if the environment in which the newly acquired "skills" are to be exercised has also changed more or less permanently. Seen in that light, Piaget's adaptation theory becomes clear. A person in development adapts slowly to the changing situation.

We may assume that in the nineties, many developmental psychologists will refer to Piaget's theory. Further elaboration and refinement of his theory will have to wait, however, until the new flood of data on the development of the processes underlying Piagetian problem-solving tasks has been integrated into a coherent new theory at a higher level.

ACKNOWLEDGMENT

This research was supported by a grant from The Open University, grant no. SW-3-96.

REFERENCES

Adey, P., & Shayer, M. (1995). *Really raising standards*. In Cognitive intervention and academic achievement. London: Routledge.

Bacon, F. (1626/1901). *New Organon I.* New York: P.F. Collier & Son.

Beilin, H. (1971). The training and the acquisition of logical operations. In M.F. Rosskopf, L.P. Steffe, & S. Taback (Eds.), *Cognitive development research and mathematical education.* Washington, DC: National Council of Teachers of Mathematics.

Bouchard, T.J., & McGue, M. (1981). Familial studies of intelligence: A review. *Science, 212*, 1055-1059.

Bruner, J.S. (1960). *The process of education.* Cambridge: Harvard University Press.

Bruner, J.S. (1966). On the conservation of liquids. In J.S. Bruner, R.R. Olver, & P.M. Greenfield (Eds.), *Studies in cognitive growth.* New York: Wiley.

Carter, L.F. (1984). The sustaining effects study of compensatory and elementary education. *Educational Researcher, 13*, 4-13.

Caruso, D.R., Taylor, J.J., & Detterman, D.K. (1982). Intelligence research and intelligent policy. In D.K. Detterman, & R.J. Sternberg (Eds.), *How and how much can intelligence be increased*? Norwood, NJ: Ablex Publishing Corporation.

Cicirelli, V. (1969). *The impact of Head Start: an evaluation of Head Start on children's cognitive and affective development.* (OEO Report No PB 184 328) Westinghoun Learning Corporation for Federal and Technical Information. U.S. Institute for Applied Technology.

Coles, M.J., & Robinson, W.D. (Eds.), (1989). *Teaching thinking. A survey of programmes in education.* Bristol: Bristol Press.

Detterman, D.K. (1982). Foreword. In D.K. Detterman & R.J. Sternberg (Eds.), *How and how much can intelligence be increased.* Norwood, NJ: Ablex Publishing Corporation.

Gal'perin, P.J. (1966). *A method, facts, and theories in the psychology of mental action and concept formation.* In Symposium 24: concept formation and inner action. 18th International Congress of Psychology, Moscow, p. 48-59.

Hooper, F.H. (1973). An evaluation of logical operations instruction in the preschool. In K. Parker (Ed.), *The preschool in action: Exploring early childhood programs.* Boston: Allyn & Bacon.

Hooper, F.H., & De Frain, J.D. (1980). On delineating distinctly Piagetian contributions to education. *Genetic Psychology Monographs, 101*, 151-181.

Hooper, F.H., & Marshall, W.H. (1968). *The initial phase of a preschool curriculum development project: final report.* Washington, DC: United States Department of Health, Education and Welfare, Office of Education.

Hunt, J.M. (1961). *Intelligence and experience.* New York: Ronald Press.

Inhelder, B., Sinclair, H., & Bovet, M. (1974). *Apprentisage et structures de la connaissance.* Paris: Presses Universitaires de France.

Kingma, J. (1981). *De ontwikkeling van quantitative en relationele begrippen bij kinderen van 4 tot 12 jaar.* [Development of quantitative and relational concepts in 4 to 12 year old children]. Lisse: Swets & Zeitlinger.

Kingma, J. (1986). The range of seriation training effects in young kindergarten children. *Contemporary Educational Psychology, 11*, 276-289.

Kingma, J., & Koops, W. (1984a). Seriatietraining bij kleuters. [Seriation training in kindergartners]. *Nederlands Tijdschrift voor de Psychologie, 39*, 234-244.

Kingma, J., & Koops, W. (1984b). Piaget en de relatie tussen onderwijs en ontwikkeling. [Piaget and the relation between education and development]. In B. Creemers (Ed.), *Onderwijskundig Lexicon.* Hertogenbosch: Samson.

Kingma, J., & Loth, F.L. (1983). Het trainen van seriatie volgens de methode van Levinova bij jonge kinderen. [Seriation training in young children according the Levinova method]. *Psychologica Belgica*, 23, 1-16.

Kingma, J., & Loth, F.L. (1984a). Effects of perceptual manipulations on number comparison tasks. *International Journal of Behavioural Development, 7*, 21-33.

Kingma, J., & Loth, F.L. (1984b). The development of relational terms in different concept areas. *The Journal of General Psychology, 111*, 163-175.

Kingma, J., & TenVergert, E.M. (1993a). Het ontwerpen van transferstudies. [Designing transfer studies]. In W. Tomic, & P. Span (Eds.), *Onderwijspsychologie: beïnvloeding, verloop en resultaten van leerprocessen.* Utrecht: Lemma.

Kingma, J., & TenVergert, E.M. (1993b). Leertheoretische en cognitivistische interpretaties van transfer. [Learning theoretical and cognitivistic interpretations of transfer]. In W. Tomic, & P. Span (Eds.), *Onderwijspsychologie: beïnvloeding, verloop en resultaten van leerprocessen.* Utrecht: Lemma.

Kingma, J., & TenVergert, E.M. (1993c). Transfer: oefening baart kunst. [Transfer: practice makes perfect]. In W. Tomic, & P. Span (Eds.), *Onderwijspsychologie: beïnvloeding, verloop en resultaten van leerprocessen.* Utrecht: Lemma.

Labinowcz, E. (1980). *The Piaget primer: Thinking, learning and teaching.* Reading, MA: Addison-Wesley.

Le Francois, G. (1984). *The lifespan.* Belmont, CA: Wadsworth.

Neimark, E. (1975). Intellectual development during adolescence. In F.D. Horowitz, (Ed.), *Review of child development research* (Vol.4, pp. 541-594). Chicago: University of Chicago Press.

Nickerson, R.S., Perkins, D.N., & Smith, E.E. (1985). *The teaching of thinking*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Nisbet, J., & McGuiness, C. (1990). Teaching thinking: The European scene. *Teacher thinking and problem solving, 12, 3,* 12-14.

Nurss, J.R., & Hodges, W.L. (1982). Early childhood development. In H.E. Mitzel (Ed.), *Encyclopedia of educational research*. New York: MacMillan.

Obuchova, L.F. (1966). *Experimental formation of representations of invariance in five- and six-year-old children*. In Symposium 24: concept formation and inner action. 18th International Congress of Psychology, Moscow. pp. 103-108.

Obuchova, L.F. (1972). *Etapy razvitija detskogo myslenija*, [Development of thinking in children]. Izdatel' Stvo Moskocvskogo. Utrecht: Rijksuniversiteit, Psychologisch Laboratorium.

Piaget, J. (1959a). *Apprentissage et connaissance. Etudes d'Epistémologie Génétique, tome 7.* Paris: Presses Universitaires de France.

Piaget, J. (1959b). Apprentissage et connaissance. *Etudes d'Epistémologie Génétique, tome 10.* Paris: Presses Universitaires de France.

Piaget, J. (1963). Le probléme de la filiation des structures. [Problem of structure association]. *Etudes d'Epistemologie Genetique, Tome 15.* Paris: Presses Universitares de France.

Piaget, J. (1964a). Development and learning. *Journal of Research in Science Teaching, 2,* 176-186.

Piaget, J. (1964b). Development and learning. In R.E. Ripple, & V.N. Rockcastle (Eds.), *Piaget rediscovered: A report of the conference on cognitive studies and curriculum development.* Ithaca, School of Education: Cornell University.

Piaget, J. (1973). La psychologie de l'intelligence. [Psychology of intelligence]. Paris: Colin.

Piaget, J. (1974). *Adaption vitale et psychologie de l'intelligence.* [Adaptation and psychology of intelligence]. Paris: Hermann.

Piaget, J (1975a). *La naissance de l'intelligence chez l'enfant.* [The emergence of intelligence in the child]. Neuchatel: Delachaux et Niesle. 8de druk (1e druk 1936).

Piaget, J. (1975b). L'Equilibration des structures cognitives: Probleme central du devolppement. [Equilibration of cognitive structures. Central problem of development]. In *Etudes d' Epistemologie Genetique* (Vol. 33). Paris: Presses universitaires de France.

Piaget, J. (1976). *Le comportement moteur de l'evolution.* [Motor behavior of evolution]. Paris: Gallimard.

Scarr, S., Weinberg, R.A., & Levine, A. (1986). *Understanding development*. New York: Harcourt Brace Jovanovich.

Schmitt, R. (1994). *Searching for Harmony*. Washington, DC: National Academy of Sciences.

Tomic, W. (1995a). Reikwijdte van transfer van een training in inductief redeneren. [Range of transfer of a training in inductive reasoning]. *Pedagogisch Tijdschrift, 20,* 341-351.

Tomic, W. (1995b). Training in inductive reasoning and problem solving. *Contemporary Educational Psychology, 20,* 483-490.

Tomic, W., Kingma, J., & TenVergert E.M. (1993). Training in measurement. *Journal of Educational Research, 86,* 340-348.

Tomic, W., & Kingma, J. (In press). On the relationship between seriation and number line comprehension.

Tomic, W., & Klauer, K.J. (1996). On the effects of an inductive reasoning training: How far does it transfer and how long do the effects persist? *European Journal of Psychology of Education*, 283-299.

Van der Zanden, J.W. (1985). *Human development*. New York: Knopf.

Vernon, P.E. (1979). *Intelligence, heredity and intelligence*. New York: Freeman.

Wadsworth, B. (1978). *Piaget's theory of cognitive development: An introduction for students of psychology and education*. New York: McKay.

Woolfolk, A.E. (1987). *Educational Psychology*. Englewood Cliffs, NJ: Prentice Hall.