

Current Topics in Human Intelligence

Volume 4

Theories of Intelligence

**Current Topics
in
Human Intelligence**

Volume 4

Theories of Intelligence

Editor

Douglas K. Detterman

Case Western Reserve University



**Ablex Publishing Corporation
Norwood, New Jersey**

Copyright © 1994 by Ablex Publishing Corporation

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise, without the permission of the publisher.

Printed in the United States of America.

ISBN: 1-56750-103-6

ISSN: 8775-0040

Ablex Publishing Corporation
355 Chestnut Street
Norwood, New Jersey 07648

Contents

- List of Contributors **vii**
- Foreword to the Series ix**
- Foreword to Volume 4 **xi**
- 1 Intelligence, Attention, and Learning: Maximal and Typical
Performance **1**
Phillip L. Ackerman
- 2 Intelligence, Personality, and Society: Constructivist versus Essentialist
Possibilities **29**
Chris Brand, Vincent Egan, and Ian Deary
- 3 Cognitive Abilities: Constructing a Theory from Data **43**
John B. Carroll
- 4 The Bio-Ecological Theory of Intelligence:
A Developmental–Contextual Perspective **65**
Stephen J. Ceci and Maggie Bruck
- 5 Theoretical Possibilities: The Relation of Human Intelligence to Basic
Cognitive Abilities **85**
Douglas K. Detterman
- 6 A Biological Theory of Intelligence **117**
H. J. Eysenck
- 7 A System for Understanding Cognitive Capabilities:
A Theory and the Evidence on Which it is Based **151**
John Horn and Jennie Noll

8	Intelligence from the Standpoint of a (Pragmatic) Behaviorist <i>Lloyd G. Humphreys</i>	205
9	Theoretical Models for the Study of Intelligence <i>Earl Hunt</i>	233
10	Phlogiston, Animal Magnetism, and Intelligence <i>Arthur R. Jensen</i>	257
11	Multiple Intelligences in Multiple Contexts <i>Mara Krechevsky and Howard Gardner</i>	285
12	CAM: A Theoretical Framework for Cognitive Abilities Measurement <i>Patrick C. Kyllonen</i>	307
13	Human Intelligence: Its Nature, Use, and Interaction with Context <i>Robert J. Sternberg</i>	361
	Author Index	409
	Subject Index	417

List of Contributors

Numbers in parentheses indicate the pages on which the author's contributions begin.

Phillip L. Ackerman, Department of Psychology, University of Minnesota, N218 Elliott Hall, 75 East River Road, Minneapolis, MN 55455 (1)

Chris Brand, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, U.K. (29)

Maggie Bruck, Department of Psychology, McGill University, Stewart Biological Sciences Building, 1205 Dr. Penfield Avenue, Montreal, QC, Canada H3A 1B1 (65)

John B. Carroll, 409 North Elliott Road, Chapel Hill, NC 27514 (43)

Stephen J. Ceci, Department of Human Development, Cornell University, Ithaca, NY 14853 (65)

Ian Deary, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, U.K. (29)

Douglas K. Detterman, Department of Psychology, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106-7123 (85)

Vincent Egan, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, U.K. (29)

H. J. Eysenck, Department of Psychology, Institute of Psychiatry, De Crespigny Park, London, England SE5 8AF (117)

Howard Gardner, Harvard Project Zero, Harvard Graduate School of Education, Longfellow Hall, Appian Way, Cambridge, MA 02138 (285)

John L. Horn, Department of Psychology, University of Southern California, Los Angeles, CA 90089-1061 (151)

Lloyd G. Humphreys, Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, IL 61820 (205)

Earl Hunt, Department of Psychology, NI-25, University of Washington, Seattle, WA 98195 (233)

viii LIST OF CONTRIBUTORS

Arthur R. Jensen, School of Education, University of California, Berkeley, CA 94720 (257)

Mara Krechevsky, Harvard Project Zero, Harvard Graduate School of Education, Longfellow Hall, Appian Way, Cambridge, MA 02138 (285)

Patrick C. Kyllonen, AL/MOAL, Brooks Air Force Base, TX 78235-5601 (307)

Jennie Noll, Department of Psychology, University of Southern California, Los Angeles, CA 90089-1061 (151)

Robert J. Sternberg, Department of Psychology, Yale University, Box 11A Yale Station, New Haven, CT 06520-7447 (361)

Current Topics in Human Intelligence
Foreword to the Series

Douglas K. Detterman, Editor

The purpose of this series is to focus on single issues of importance to the study of human intelligence. Unlike many edited volumes, this one is designed to be thematic. Each volume will present a detailed examination of some question relevant to human intelligence and, more generally, individual differences.

The reason for beginning this series is that a forum is needed for extensive discussion of pertinent questions. No such forum currently exists. A journal does not allow an author sufficient space or latitude to present fully elaborated ideas. Currently existing edited series are not thematic but, instead, offer researchers an opportunity to present an integrative summary of their own work. Both journals and edited, nonthematic monographs are essential to the advancement of the study of human intelligence. But they do not allow collective intelligence to be brought to bear on a single issue of importance.

Why is it important to examine specific issues in detail? The answer to this question depends on an appreciation of the historical development of the study of human intelligence. For at least 40 years, and perhaps longer, the study of human intelligence has been less than highly regarded as an academic pursuit. The reasons for this attitude are many and have been discussed elsewhere. Despite the reasons, the lack of academic sanction for the study of individual differences in human intelligence has produced a discipline without a unifying paradigm. When researchers study human intelligence, it is almost always from the perspective of training in a related, but different, discipline. Disciplines that have “sacrificed” researchers to human intelligence include: cognitive,

developmental, and educational psychology, behavior genetics, psychometrics, mental retardation, neuropsychology, and even experimental psychology. Each of these researchers has brought a different set of assumptions and methods to the study of human intelligence.

That so many different points of view have been applied to a single subject area has, in my opinion, brought vitality to the endeavor, a vitality currently lacking in so many areas of the social sciences. But this vitality does not arise from the isolation and fractionation that can be the result of different points of view. On the contrary, it results from the juxtaposition of these different points of view, a juxtapositioning that has occurred with increasing frequency over the last 10 years.

Therefore, it is the purpose of this series to bring different points of view together on issues of importance to understanding human intelligence. The hope is that, at the very least, researchers will find more reason to give their primary allegiance to the study of human intelligence and, at the most, the series will contribute to the emergence of a unifying paradigm.

Foreword to Volume 4: Theories of Intelligence

There is no doubt that the area of human intelligence has generated more empirical data than any other area in the social sciences. There are literally thousands of studies that relate the score on an IQ test to almost any other biological, psychological, or social characteristic that exists. We know the relationship of height, tooth decay, brain size, physical attractiveness, and even penis size to IQ. It is difficult to think of anything that has not been correlated with IQ.

With this rich empirical foundation, there should be more theories of intelligence. In many areas of psychology, theories are more prevalent than data. I recently heard someone doing memory research talking about toothbrush theories. Memory theories are like a toothbrush because everyone has one and no one uses another person's. This is not true of intelligence. Though many people propose definitions of intelligence, few formalize their speculation in a theory.

That is what this book is about. Many of the best researchers in the field were asked to present their theory of intelligence. The resulting chapters are presented in alphabetical order by first author. The reason for this organization has nothing to do with fairness: I couldn't find an organizing principle that would allow me to sort the chapters into broad topic areas. What characterizes these chapters is their diversity. It would be an interesting exercise as you read this book to keep a list of the disputed theoretical issues—it would be lengthy.

A more important list would be the number of issues on which this wide assortment of authors agree. Despite the diversity of viewpoints, all or nearly

all of the authors agree that certain phenomena have to be explained by any successful theory of intelligence. I think that it is an important strength of the study of intelligence that there is agreement on what the important issues are. The list of issues that are agreed on is longer than those that are not agreed on. Authors may not always agree on the resolution of these issues but they do agree that they are important.

The importance of having theories of intelligence that can be tested cannot be underemphasized. One function of a theory is to summarize the empirical facts about a phenomenon. Given the large number of studies of intelligence, summarization is badly needed. Theories also can provide guidance about what research should be done next and focus the mind on important issues to make us think of the bigger picture. Such guidance is always welcome, but it is particularly important at a time when new methodologies are opening a window on the brain in action.

I hope the chapters in this book will begin an ongoing debate that will result in a good theory of human intelligence that is widely accepted.

Douglas K. Detterman
July 24, 1993

Chapter 1

*Intelligence, Attention,
and Learning:
Maximal and Typical
Performance**

Phillip L. Ackerman

*Department of Psychology
University of Minnesota*

*Preparation of this chapter was made possible through the support of the Office of Naval Research, Cognitive Science Program—Contract N00014-89-J-1974, Phillip L. Ackerman, principal investigator.

INTRODUCTION

Overview

In advancing a new theoretical perspective for intelligence assessment and application, this chapter focuses on two major themes. The first theme concerns the potential misspecification of intelligence as a construct solely delimited by maximal performance. The original assumptions regarding the desired conditions for testing and the inferences to be made regarding the construct of intelligence will be challenged, and an alternate view that includes “typical” performance will be proposed. Even though the concept of “typical” performance requires additional specifications for (a) volitional, (b) affective, and (c) dynamic constructs, it is proposed here that a theoretical specification of intelligence from such a perspective will be both more accurate (in terms of generalizability to the “real world”) and effective (in terms of predictive/concurrent validation strategies).

The second theme of this chapter concerns specification of a theory of intelligence, and of the nomological network that relates constructs of intelligence, attention, and learning. It will be argued that, for many intents and purposes, it is heuristically useful to consider intelligence (and its inherent notions of interindividual differences) as equivalent to the construct of attention. A comparison of respective structures for intellectual abilities and attention will be discussed, and parallels will be drawn and evaluated. Following discussion of the justification for the basic premise of construct overlap between intelligence and attention, it will be demonstrated that the integrated perspective affords a number of predictions concerning the expression of intelligence in the processes and products of learning situations. Results from several empirical studies will be discussed from the perspective of this construct integration.

Given the limited nature of research directly bearing on the unified theory developed here, the final section of this chapter is devoted to specification of tactics and strategies for extension, application, and evaluation of the proposed theory.

Specification of the Proposed Theory

There are four principal tenets to the proposed theory of intelligence, as follows:

1. *Maximal vs. typical performance.* The psychometric premise that intelligence is best associated with maximal performance is incomplete. A more comprehensive view recognizes that the quantity and quality of “intelli-

gence engaged” has properties of “typical” as well as “maximal” performance.

2. *Intelligence and attention.* The engagement of intelligence is described fundamentally by the operations of attentional mechanisms. Such engagement is limited for individuals, both extensively (i.e., amount of attention engaged) and temporally.
3. *Intelligence and learning.* The initial phases of learning (especially in terms of development of automatized knowledge—whether procedural or declarative) is a fundamental application of intellectual engagement.
4. *A unified perspective.* Performance on many so-called “intellectual” tasks is critically dependent on (a) level of potential (or maximal) intellectual/attentional engagement, (b) basal level (or typical) intellectual/attentional engagement, (c) self-regulatory/motivational processes that determine the *recognition* of intellectual/attentional demands, and the (d) *decisional processes* that bring about responses to perceived intellectual/attentional demands.

Each of these tenets of the theory is discussed in turn.

MAXIMAL VS. TYPICAL INTELLIGENCE

Cronbach (1949) first created a scheme that classified ability tests as measures of maximal performance and personality tests as measures of typical performance. In a subsequent series of papers, Fiske and Butler (1963, Butler & Fiske, 1955) explicated this contrast between ability testing (“Try to do your best”; Fiske & Butler, 1963, p. 251) from personality testing (“We are ordinarily concerned with the typical (modal or mean?) strength of this tendency [to respond in a given way] because this provides the best estimate of what a person is most likely to do”; p. 258). The authors presented a two-fold rationale for assessing ability at its “maximum” level:

First, we want a pure measure, one that is determined almost wholly by one thing, the subject’s capacity, rather than a measure which is affected by several influences. Second, we measure maximum performance because it is probably more stable than performance under more lifelike conditions. (Fiske & Butler, 1963, p. 253)

A main purpose of this chapter is to question these stated reasons for focusing on maximal performance. By rejecting these fundamental assumptions about the testing situation, a more complex view of intelligence will be proposed, one that is more comprehensive and yet applicable to many situations where traditional measures of intelligence have lacked validity.

Capacity

Many intelligence theorists have discussed whether extant tests actually focus on ability-as-capacity or ability-as-developed trait (e.g., Humphreys, 1979; Wechsler, 1944). With the notable exception of the fluid intelligence construct proposed by Cattell and Horn (Cattell, 1963; Horn & Cattell, 1966), *most* modern theorists and psychometrists find notions of capacity measures cause for considerable uneasiness. The main reasons for discomfort with ability-as-capacity assessment are as follows.

Developmental/experiential/cultural components to abilities. From the earliest efforts of Binet and Simon (1905), psychometricians have made many attempts to *minimize* the impact of cultural and environment-specific influences on intelligence performance. However, although a goal of determining intelligence through testing *independent* of prior history can be seen as admissible, it is recognized that experiential effects cannot be partialled out of the equation (nor, indeed, should they be). Part and parcel of any practically useful test of intelligence is the test's efficiency in predicting aspects of current or future performance; performances which are undeniably influenced by the individual in the context of interaction with his/her environment.

The concept of capacity is, in a lexical sense, a concept of permanence. However, a central finding in the assessment of intelligence is that changes in test performance occur as a result of maturation and aging (from childhood to adulthood, or even across the adult life span) (e.g., see Bayley, 1949, 1968; Honzik, MacFarlane, & Allen, 1948). Moreover, experiential differences across individuals (e.g., schooling, life experiences) often result in other changes to test performance (see Flory, 1940; Husen, 1951). With such a database to address, it makes little sense to refer to *unstable capacities* of the individual. It has been repeatedly shown (e.g., Thorndike, 1940) that intelligence test performance changes over time to a greater degree than would be expected through measurement error alone. Similarly, to identify an individual's "capacity" as their maximal performance on a particular test would require specification of what experiential or maturational conditions would have had to be satisfied in order to match capacity with performance. In addition, given some early deprivations, it may not be possible, even in theory, to specify the conditions that would allow an individual to develop performance to a level that would approximate his/her theoretical capacity (i.e., if an optimal environment had been available throughout the lifespan). Clearly, as many modern theorists have maintained, intelligence is much more parsimoniously thought of in terms of current level of functioning, either on an absolute scale (e.g., for memory span) or on a relative scale (e.g., the Deviation IQ).

Ability cannot be divorced from motivation and/or personality. The instructions to test administrators often emphasize the importance of adequate rapport between the administrator and the test taker. Such an emphasis is

explicit acknowledgment (supported by some empirical investigations, e.g., Gordon & Durea, 1948; see review by Sattler & Theye, 1967) that the motivation for performing well on an intelligence test is a critical ingredient for accurate measurement of maximal performance. However, as numerous investigators have demonstrated, both the physical state of the test taker (e.g., time of day, arousal level; see Revelle, Humphreys, Simon, & Gilliland, 1980) and the psychological state of the test taker (presence or absence of debilitating anxiety, incentives, motivation to succeed or fail, degree of interest in the testing situation) may significantly impact the level of test performance (Quereshi, 1960). It thus makes little sense to discuss ability-as-capacity, unless the concept of capacity is substantially subverted to allow for changes over the course of the day or in line with the individual's volitional state.

The interdependence of ability and volition constructs have been recently noted by Cronbach (1990):

Typical response and ability are not truly separable. A person's record on typing tests establishes that his ability has reached some level, but the score also reflects his willingness to push himself in that kind of situation. That is to say, [the] distinction does not take purposes into account. Purposes are important, but they have been given little consideration in the psychology of individual differences. . . . Test directions almost never tell the examinee how to approach the task. Theoretically, style of performance falls under the head of typical response; but variation in style affect "ability" scores. (p. 38)

This point is concordant with a broader context of intelligence-as-typical performance.

Stability

Fiske and Butler (1963) argued that measures of maximal performance are apt to be more stable than measures of typical performance. This proposal was predicated at least partly on the basis of their contrast between the stability of personality measures (described as measures of typicality) and the stability of ability measures (which they termed measures of capacity). However, given that personality and ability measures differ on many other dimensions (e.g., ability test items have *correct answers* answers and personality test items do not), it has never been clear that the stability of ability tests is *caused by* conditions designed to obtain maximal performance. Or, rather, that relative stability is caused by other sources of variance, or by inherent differences in trait stability, not the least of which are the controversial aspects of situational specificity of personality. It simply has not been established that, when instructed to "do one's best" on a personality test (as might be told to an individual who is taking the test under job screening circumstances), personality

test scores become more or less stable. Nor, has it been established that instructions to examinees to “respond as you typically would; say, if this was a puzzle in a magazine” to an ability test, make ability test scores less stable.

It seems no less desirable to a psychology of intelligence than it does to a psychology of personality that a construct measure provide the best “estimate of what a person is most likely to do” (Fiske & Butler, 1963, p. 259). Even if future research were to show that measures of intelligence-as-maximal performance turn out to be more stable than measures of intelligence-as-typical performance, one must also keep in mind that there is a potential for simultaneously *increasing* reliability and *decreasing* validity (e.g., such as when test items are made more homogeneous).

To date, there apparently have been no empirical data reported to support Fiske and Butler’s conjecture that tests of “maximal” performance are more stable than tests of “typical” performance when a single construct is considered. Moreover, given the argument that a desire for stability is not necessarily a vector model (in preference scaling, a vector model corresponds to “more is always better”), it follows that Fiske and Butler’s arguments about the superiority of maximal performance measures for the sake of stability are spurious ones for defending the sole examination of maximal measures of ability.

Cognitive, Information-Processing Approaches to Intelligence

The same assumptions made by Fiske and Butler (1963) regarding the circumstances for ability testing are ubiquitous in studies of the information-processing components of intellectual abilities. In fact, while most of the time such an assumption of maximal performance is implicit to the experiment (usually through instructions to subjects to “do your best” or “work as fast as you can, without making errors”), often the assumption is explicit (as when subjects are excluded from analysis because they didn’t “try hard” on the task). Because most studies of information processing and intellectual abilities test both information processing and abilities in the same laboratory context, they represent a special case of the states of the world (i.e., when motivation to perform on the information-processing task is likely to be concordant with motivation to perform on the ability tests, given under identical instructions). A similar argument for examining information processing under “typical” circumstances seems appropriate. It might not be as far-fetched as it seems on the surface to explain what kinds of information processing people “actually do,” rather than only what they “can do.” Some inroads have been made in this domain, by investigations of reasoning and problem solving in everyday contexts (e.g., Sternberg & Wagner, 1989), and in the context of research on human error (e.g., Reason, 1990; Reason & Mycielska, 1982), but much more work remains to be done in constructing ecologically valid theories of information processing.

Summary

The central point of these arguments is that the construct of intelligence has become overly restrictive in the 90 or so years of modern theory and research. By expanding the construct to include a dimension bounded by maximal and typical levels of intellectual engagement, it is proposed that a better understanding can be derived for how intellectual abilities interact with other personal constructs (e.g., motivation and volition). In order to generalize to the world outside the testing room or the laboratory, intellectual abilities must be placed in the context of everyday life.

By considering typical performance and maximal performance aspects of intelligence in light of attention and learning issues, a more comprehensive view may be obtained that has the potential to improve prospects of predictive validity for school and job performance, to predict developmental changes in ability-as-maximal performance, and to integrate ability with personality and volitional determinants of performance. The topics of attention and learning will be discussed first, and then implications of the unified approach to intelligence will be presented.

INTELLIGENCE AND ATTENTION

Several years ago, in attempts to establish linkage between intellectual abilities and information-processing concepts of the learning process, I was struck with the prominent role *attention* is accorded in normative cognitive research. Attention is a critical determinant of the acquisition of knowledge and skills, performance of multiple tasks, performance of tasks with novel or inconsistent information-processing requirements, and so on (e.g., see Ackerman, 1986; see also Zeaman, 1978, for an earlier discussion relating intelligence and selective attention). In addition, it was clear from the literature that attentional requirements of tasks are *not* static, and that predictable changes of attentional demands take place as an individual acquires knowledge or develops information-processing skills.

The historical treatment of the structures of the constructs of the intelligence and attention (Figure 1) provides some interesting parallels. To retrace some familiar territory, an abbreviated history of formal theories of intelligence always includes Spearman's theory. Spearman (1927) envisioned intelligence to be primarily unitary (or *g*), with each ability test being partitioned into a general intelligence component, and a component specific to that particular test. Although the formal theorizing about attention has only begun in more recent years, Kahneman's (1973) view of attention was quite similar to Spearman's theory of intelligence. Kahneman proposed a single, undifferentiated source of attentional capacity (denoted in the figure as "A"). In addition, Kahneman posited that spe-

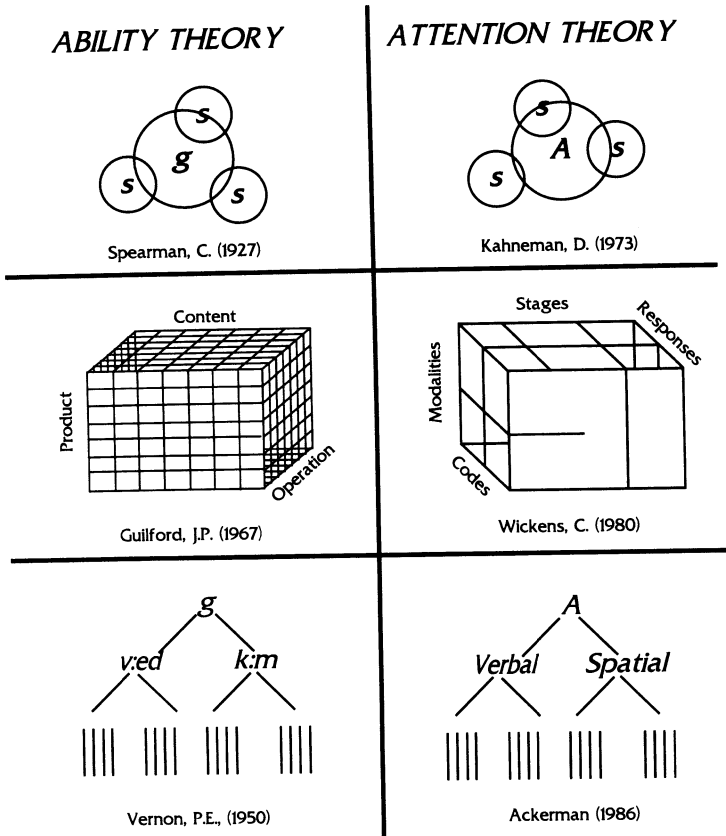


Figure 1. An abbreviated historical comparison of structural representations for intelligence theory with attention theory. Top pair of structures have a general ability/process as central, middle pair of structures are orthogonal processes, with no general process, and the lower pair of structures are hierarchical, with a general component, but also with subordinate, more specific components. The set-theoretic illustration of Kahneman's (1973) theory is drawn as inferred, and has not been previously represented in this fashion.

cific structures (such as the visual, auditory, and tactual systems) were important, and that these structures were responsible for performance conflict among some varieties of dual-task and multiple-task situations. From an abstract point of view, though, both Spearman and Kahneman viewed these respective constructs of intelligence and attention as a single general process or system.

Guilford's (1967) extension of Thurstone's theory of intelligence serves as the next landmark in this comparison among intelligence and attention theories, and functions as contrast to Spearman's theory. The three-dimensional

Structure of Intellect Model derived by Guilford designates *independent* sources of intellectual operations (i.e., Contents, Operations, and Products). On the attention side, we have Wickens's (1980) Structure of (attentional) Resources Model. These two box models show a surface-level similarity. More importantly, though, the models are similar from an underlying theoretical perspective. Wickens holds, for instance, that there is no general source of attention, but there *are* separate, independent Spatial and Verbal pools of attentional resources, just as the Guilford model proposes that analogously separate and independent Symbolic and Semantic contents.

Finally, most current views of intelligence are generally in agreement with a hierarchical structure of intellectual abilities, such as the one by Vernon (1950) depicted in Figure 1. Contemporary views maintain the importance of some general intellectual factor, but acknowledge that other intellectual abilities, with varying degrees of generality-specificity coalesce from the patterns of individual differences on the universe of tests. This view has not been incorporated into the attention field, perhaps because it is early for such a reconciliation to take place. However, both recent and earlier experiments do point to the fact that both the Kahneman and Wickens models have merit. Thus, the sketch of a hypothetical model of attention and I have outlined here, which incorporates the two models, does not go beyond the existing data—only beyond current theory.

With this background in mind, we can consider the construct of general or broad-content forms of intelligence as either a maximal value (associated with allocation of all of one's attentional effort) or as something less than maximal (associated with allocations of some portion of the total attentional effort possible). As such, individuals may differ on the maximal amount of attentional/intellectual effort available to them, and to the proportion of their own attentional/intellectual effort actually devoted to a task. From a broad perspective, effort devoted to a task, then, is the product of the individual's total effort available and the proportion of available effort devoted to a task.

Extensive and Temporal Aspects of Attentional Resource Availability

It is a profoundly erroneous truism, repeated by all copy-books and by eminent people when they are making speeches, that we should cultivate the habit of thinking of what we are doing. The precise opposite is the case. Civilization advances by extending the number of important operations which we can perform without thinking about them. Operations of thought are like cavalry charges in battle—they are strictly limited in number, they require fresh horses, and must only be made at decisive moments. (Whitehead, 1911, pp. 41–42)

Although there has been little work examining the short-term (e.g., minutes and hours) course of available intellectual engagement, there are numerous examples of the limits of attentional capabilities over time, in terms of the

learning environment (e.g., see Zeaman & Kaufman, 1955) or sustained attention (e.g., vigilance research; see Parasuraman, 1984, for a review). Depending on the experimental paradigm under consideration, it appears that attention is only sustainable for periods of about 30 minutes before substantial performance decrements occur. Changes in the task environment often have the effect of attenuating such decrements, but it remains clear to even the most causal observer (as it was to Whitehead, in the quotation above), that there are costs associated with sustained engagement of attentional effort.

Intelligence researchers typically blanch at the idea of giving subjects a continuous set of tests for an 8- to 12-hour period. Even though there have been no parametric studies of testing time limits that I am aware of, many experiments reported in the literature limit testing to a few hours at a time. As such, there seems to be an implicit acknowledgment that it is impractical to ask subjects to maintain their intelligence-as-maximal performance levels for extended periods of time.

In order to equate intellectual and attentional engagement, it is critical to examine the determinants of attentional investment during maximal engagement and typical engagement, and the course of attentional depletion over extended periods of effort. Certainly, research is needed to determine the time course of intellectual/attentional engagement, but the important consideration again, is that intelligence-as-maximal performance may have substantial limits in generalizability to the external world, give that an individual's engagement with the real world runs the course of the entire waking day, and not just a few hours in the laboratory or the testing room.

INTELLIGENCE AND LEARNING

The starting point for consideration about intelligence and learning is with E. L. Thorndike (in Thorndike, Bregman, Cobb, Woodyard, et al., 1926). Although much of Thorndike's work was cast in the associationist framework, he captured the importance of *learning* as a critical element of intelligence, stating: "An obvious hypothesis, often advanced, is that intellect is the ability to learn, and that our estimates of it are or should be estimates of ability to learn" (p. 17). Although Binet's approach to intelligence can be thought of as focusing on learning in the broader context of educational achievement, Thorndike was instrumental in bringing a microlevel orientation to learning—one that could be mapped to then-current experimental inquiry. Many investigators followed Thorndike's lead in focusing on the relationship between intelligence and learning; most notable of these research programs are those by Woodrow (e.g., 1946), Ferguson (e.g., 1954, 1956), and Cronbach and Snow (e.g., 1977). Reviews of salient contributions from these investigators regarding the relationship between intelligence and learning, along with discussion of

methodological and statistical problems associated with assessing such relations, lie beyond the scope of this chapter. However, readers who wish to consider extant reviews of these developments can consult articles by Adams (1987) and by Ackerman (1987).

Several fundamental aspects of the relationship between the constructs of intelligence and learning serve as the foundation for the current theoretical formulation. These include (a) intellectual abilities and skill acquisition, (b) intelligence and academic learning, (c) dynamic effects of ability and volition, and (d) other aptitude-treatment interactions. Each is discussed briefly next.

Intellectual Abilities and Skill Acquisition

Through a series of empirical and theoretical investigations, I have previously argued for a dynamic set of relations between intellectual and individual differences in task performance during three phases of skill acquisition (e.g., see Ackerman, 1988; for a review of the phases of skill acquisition, see Fitts & Posner, 1967). The relations between ability and performance during skill acquisition are illustrated in Figure 2. During the initial, cognitive phase of skill acquisition, the critical elements of intelligence-as-maximal performance (general intelligence and broad content factors of ability) are the most important determinants of individual differences in task performance. As practice proceeds on a task that has consistent information-processing demands (the associative phase of skill acquisition), the role of general and broad content abilities diminish, while measures of perceptual speed increase in their power of predicting individual differences in task performance. Finally, at asymptotic autonomous (or proceduralized) stages of skilled performance, the role of perceptual speed abilities attenuate, in favor of associations between psychomotor abilities and task performance.

The validity of this theory has been demonstrated for a variety of tasks, including Sternberg-type memory/visual search tasks, choice reaction-time tasks (Ackerman, 1988), and complex air traffic controller simulation tasks (Ackerman, 1988; Kanfer & Ackerman, 1989). In addition, recent extensions of this theory have shown a generalization to the dynamic characteristics of perceptual speed and psychomotor tests, when given under practice conditions (Ackerman, 1990). From these various demonstrations, general and broad content (spatial, verbal) abilities are most highly correlated with initial performance on consistent tasks, perceptual speed abilities are most highly correlated with performance at intermediate stages of skill acquisition, and psychomotor abilities are most highly correlated with performance at late, autonomous stages of procedural skill performance.

Much of the representation of early portions of the learning curve is concordant with the theory and empirical research developed by members of the Air Force Human Resources Laboratory-Project LAMP group. Although there

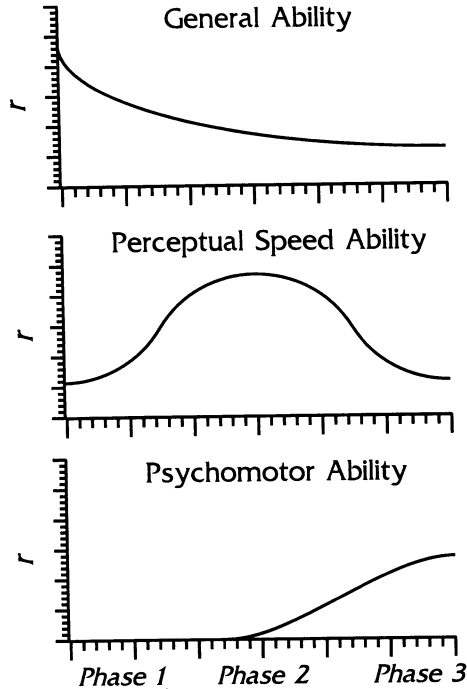


Figure 2. Hypothetical ability-performance correlations during three phases of skill acquisition, as proposed by Ackerman (1988). The hypothetical task is moderately complex, and provides for consistent information processing. From "Determinants of Individual Differences During Skill Acquisition: Cognitive Abilities and Information Processing," by P. L. Ackerman, 1988, *Journal of Experimental Psychology*, 117.

are some salient differences between the starting points and research paradigms I and the Project LAMP group have used, there is substantial agreement regarding the intense demands the tasks place on the learner's attention when they are initially encountered (e.g., see Kyllonen & Christal, 1990; Kyllonen & Woltz, 1989; Woltz, 1988). Woltz (1988) demonstrated how the "intensive" aspects of working memory (which are analogous to attention) are important in initial performance of skill acquisition tasks, most notably when formulation and manipulation of declarative knowledge demands are highest, while "activation savings" aspects of working memory, which are a representation of memory efficiency, are more highly associated with later, proceduralized knowledge. Kyllonen (in Ackerman & Kyllonen, 1991) has further argued that the subjective mental workload for tests of working memory capacity is substantially higher than for tests of existing knowledge. In a further series of con-

struct validity studies, Kyllonen and Christal (1990) demonstrated substantial (if not complete) overlap between the intensive aspects of working memory and reasoning ability. These findings can be considered as substantial support for the construct overlap between attention and general intelligence.

From the current perspective, the role of intelligence-as-maximal performance is most closely associated with the initial stages of skill acquisition, while it is posited that intelligence-as-typical performance is conceptually more associated with (a) further gains in task performance beyond the cognitive phase of skill acquisition, and with (b) asymptotic or proceduralized skilled performance. This framework is consistent with the notion that initial confrontation with a novel task involves attentional *demands* on the learner that are typically hard to ignore. Thus, even with only modest motivation for performing well, the learner is spurred to engage most or all of his or her intellectual/attentional capabilities. However, once the learner arrives at a level of performance that meets that basic task demands (i.e., attaining an acceptable level of accuracy or speed on the task), those intellectual/attentional demands are diminished, and the task moves from being predominantly attentional resource-limited to predominantly attentional resource-insensitive (Norman & Bobrow, 1975). As Figure 3 illustrates, substantial reductions of intellectual/attentional engagement, at an intermediate stage of skill acquisition, no longer have the dire consequences to performance that are found in the early resource-limited stage of skill acquisition. Therefore, when the *demands* of the task are reduced, those factors that determine intelligence-as-typical performance (such as personality or motivational factors) increase in influence. When there is a mismatch between intelligence-as-typical performance and intelligence-as-maximal performance, the learner may be satisfied with a suboptimal level of task performance, and essentially “exit” form the learning process.

Laboratory research regarding the effects of goal setting (a motivational/self-regulatory intervention) on task performance at varying stages of skill acquisition supports the current perspective. In a series of studies by Kanfer and Ackerman (1989), a goal-setting manipulation, when given at an early stage of task engagement, had no positive effect on performance. Indeed, when task attentional demands were high, the goal-setting manipulation had an overall negative impact on performance, which was accentuated for those learners of lower intellectual ability levels. However, once the declarative knowledge (or cognitive) phase of skill acquisition had been completed, the goal-setting manipulation brought about significant gains in performance, especially for those learners who were found to have lower levels of intelligence-as-maximal performance. As such, the role of these motivational mechanisms became a more important determinant of task performance during the later phase of skill acquisition, *even as correlations between general intelligence and task performance declined.*

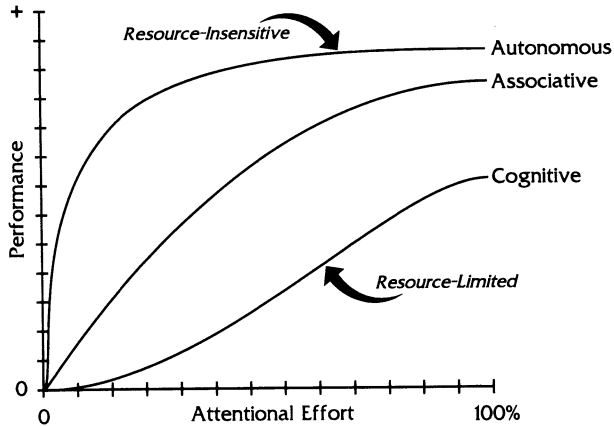


Figure 3. Hypothetical performance-resource functions for a task at three stages of practice: Cognitive, Associative, and Autonomous. Cognitive phase is predominantly resource-dependent and Autonomous phase is mostly resource-insensitive.

Intelligence and Academic Learning

Another data set, albeit a strictly correlational one outside the laboratory, cogently argues for the dynamic relations of intelligence and learning. The study by Lin and Humphreys (1977) gives perhaps the clearest illustration of the dynamic effects of intelligence-as-maximal performance and intelligence-as-typical performance with respect to academic achievement. In this study, the Graduate Records Exam (GRE) was used in predicting and postdicting graduate school and undergraduate grade point averages (GPA), respectively. GRE data were collected on several hundred students, with the exam typically taken in the senior (fourth) year of college. Correlations between GRE scores and the GPA were calculated for each of the four years of undergraduate education and six semesters of graduate school (in the fields of physics, chemistry, and mathematics). The relevant data are depicted in Figure 4.

There are two salient features of the data plotted in the figure. First, the magnitude of test validity was *not* merely a decreasing function of temporal displacement between time of test and time of GPA assessment. That is, this particular aspect of the data contradicts the notion (Henry & Hulin, 1987; Hulin, Henry, & Noon, 1990) that all test validity coefficients must decline with increasing temporal displacement between time of test and time of criterion measurement. The second feature of the figure, and the one that directly fits with the current proposal regarding the relations between intelligence and learning, is that the highest correlations between the GRE (arguably a measure

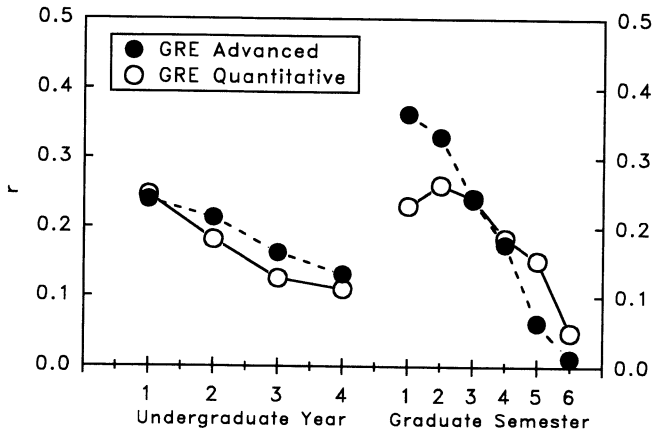


Figure 4. Validity coefficients for GRE Advanced and Quantitative tests for graduate students in math, chemistry, and physics, over four undergraduate years and six graduate semesters. *Postdiction* of undergraduate grade point average; *prediction* of graduate grade point averages. From data reported by Lin and Humphreys (1977).

of general intelligence-as-maximal performance) and GPA (a measure of academic achievement or learning), occur for the initial confrontation with both the undergraduate (Year 1) and graduate (Semester 1) systems. These situations are quite aptly described as the earliest phases of learning, when students must develop new strategies to perform their tasks. Later phases of learning, after initial confrontation with the novel environments, show declining validities associated with intelligence-as-maximal performance.

Dynamic Effects of Ability and Volition

An additional data set comes from a study by Helmreich, Sawin, and Carsrud (1986) of the relations between motivational/personality dispositions and job performance, from early job performance to performance after 6–8 months of job tenure. The sample studied included 268 telephone ticket agents working for an airline. Three samples of performance were taken, namely average job performance for 1–3, 4–6, and 7–8 months on the job. Although no ability data were reported in that study, the authors reported that their measure of achievement motivation (called “work orientation”) is essentially uncorrelated with “typical” measures of intelligence.

The theory under consideration here proposes that motivational dimensions are less important determinants of performance than intelligence-as-maximal performance early in skill acquisition, but potentially more important when skills have been learned (and both intelligence-as-typical performance and its volition-

al correlates *are* important determinants of performance). Consistent with the current proposal, the correlations between work orientation and job performance were $r = .09, .336,$ and $.34,$ for the three periods of performance sampling, respectively. That is, the dimension of achievement motivation started off as an unimportant determinant of initial job performance, but became a salient (and significant) predictor of performance after substantial job experience.

Other Aptitude-Treatment Interactions (ATI)

In addition to the fluctuations in motivational mechanisms that result from changes in intentions or incentives, other more enduring personal properties of individuals appear to have substantial impact on intelligence-as-typical performance. Individual differences in personality dimensions, such as anxiety, achievement via independence or achievement via conformance (see Cronbach & Snow, 1977; Snow, 1989a, 1989b), and introversion/extraversion (Lynn & Gordon, 1961; Revelle, 1989) have also been implicated in determining the nature of typical performance on intellectual tasks and tests. Indeed, one common assumption regarding ATIs in the academic domain is that maximal performance is elicited under *different* treatments for learners of differing aptitude or trait profiles. Although a review of these areas is beyond the scope of this chapter (see Snow & Yalow, 1982), it seems that an intelligence-as-typical performance construct may be used to clarify many of the existing ATIs, as well as provide the basis for higher order interactions, that is, those which include an interaction with the discrepancy between intelligence-as-maximal performance and intelligence-as-typical performance for each individual.

A UNIFIED PERSPECTIVE

Determinants of Performance

The framework proposed here indicates that optimal prediction of performance on any of a wide variety of cognitive tasks requires knowledge of four critical determinants of performance:

1. *Level of intelligence-as-maximal performance.* This construct corresponds to the ideal of current intelligence testing paradigms. Intelligence-as-maximal performance is the level of performance by an individual that is obtained under the conditions that are optimal for drawing out that individual's best efforts, while minimizing the deleterious effects of distractions. (For example, in Atkinson's theory [1974], this would correspond to moderate levels of arousal; in Revelle's [1989] theory, it would correspond to different levels of induced arousal for introverts and extroverts.) To date, this appears to be obtained in

testing conditions where examinees are highly desirous of doing well, and test time is limited to a few hours.

2. *Basal (typical) level of intellectual engagement.* This construct corresponds to the usual level of attentional/intellectual effort put forth in the normal day-to-day environment of the individual. Essentially, intelligence-as-typical can be thought of as an individual's average level of effort elicited across situations as varied as school, work, recreation, family interactions, and so on. It is expected that there are personality, volitional, and interest correlates of intelligence-as-typical performance (e.g., see Dreger, 1968; Humphreys & Revelle, 1984; Kanfer, 1987).

3. *Self-regulatory skill level.* Self-regulation, especially in the form of self-monitoring, appears to determine the *recognition* of intellectual demands, and thus sets the stage for decisions made regarding the level of effort allocated to a task (below) (e.g., see Kanfer, 1987; Kanfer & Ackerman, 1989). Individuals who are low in self-monitoring skills are expected to have a generally lower *effort demanded–effort supplied* coherence, whereas individuals who are high in self-monitoring skills are expected to have a higher *effort demanded–effort supplied* coherence. That is, effective self-monitoring brings about accuracy in evaluating attentional/intellectual demands of tasks. Increasing accuracy in these demand evaluations can be thought of as doubly important—because the individual will be able to allocate effort when performance is resource-dependent, *and* because the individual will be able to conserve resources over time, when the task is resource-insensitive, so that resources will not be depleted when they are called for during other activities.

4. *Decisional processes and dispositional tendencies that bring about responses to perceived intellectual demands.* An individual's recognition of intellectual demands is just one element in the determination of how much attentional/intellectual effort that an individual devotes to a particular task. Other decisional processes that are made up of metacognitive and metamotivational strategies help determine how effort is allocated across various demands (e.g., on-task and off-task effort), see Kanfer (1990b) for a review. Similarly, dispositional tendencies toward intellectual tasks (e.g., achievement motivation, or learning/performance orientation) affect the valences of particular tasks. Although these later processes are more properly considered motivational and/or affective, they play an important mediating role in the engagement of intellectual effort (Kanfer, 1990b; Kuhl, 1981; Revelle, 1989).

Toward Assessment of Typical Intelligence

It is proposed here that the construct of intelligence-as-typical performance must supplement the construct of intelligence-as-maximal performance. There are several key reasons for this particular proposal:

1. Predictions of typical or everyday intellectual performance, especially long-term performance, based on measures of maximal performance, result in a construct mismatch. The employment or school system domains cannot normally be expected to impose the same contingencies as are found in the usual selection test environment, or the environment where scholastic aptitude measures are administered. Once the job applicant is hired, or the school applicant matriculates, how the individual is *likely* to perform is best matched by that individual's intelligence-as-typical performance.

2. Generalization beyond the classroom or the job is possible. One of the most enduring criticisms of modern intelligence measurement is that it revolves around items that are relevant only to academic performance (and by implication or explicitly, not to everyday life). The construct of intelligence-as-typical performance appears, at least on the surface, to be able to encompass many aspects of intelligence that fall outside the traditional academic domain (e.g., see Sternberg, Conway, Ketron, & Bernstein, 1981, for a description of lay concepts of intelligence). Measures derived from the intelligence-as-typical performance perspective proposed here might well include items that concern such things as pursuit of knowledge and skills outside the classroom (e.g., reading for pleasure, challenging oneself to learn new skills, keep up with current events, and so on). In addition, such measures of typical intellectual engagement may be more highly predictive of developmental changes in intelligence-as-maximal performance, in that the intelligence-as-typical performance measures may capture the essence of the individual's ongoing engagement with the world at large.

3. When coupled with measures of intelligence-as-maximal performance, measures of intelligence-as-typical performance may provide more accurate predictions of school or job performance at all levels of the respective systems. That is, the fault of typical SAT/GRE exams is that they, as measures of maximal performance, are only good predictors of initial performance (in school, or in an employment training program), see Lin and Humphreys (1977) and Humphreys and Taber (1973) for examples. However, measures of intelligence-as-typical performance may provide incremental validity for prediction of performance when the demands for maximal intelligence are attenuated. Placing maximal and typical performance together in a multiple regression equation may make it possible to accurately predict performance at early and late stages of knowledge/skill acquisition.

4. A forum for examination of ability/motivation/personality interrelations. Many extant theories of personality and motivation have minimized the relations between these constructs and intellectual abilities. However, with perhaps the exception of some goal-setting studies and some arousal studies (e.g., see Kanfer, 1990b, for a review of goal-setting; and Revelle, 1989, for a review of arousal and personality), there have been mismatches between the situations for study of personality and motivation, and situations for assessment of intel-

lectual abilities. Given the potential implausibility of personality assessment under “do your best” instructions, discovery of overlap among constructs of intelligence and personality seem most likely under assessment under “typical” conditions for both.

5. An opportunity to evaluate the notion that there is a “recognition” or metacognitive component to the typical engagement of intelligence. The present contrast between maximal and typical aspects of intelligence implicitly assumes that intelligence does not function like an engine that is always running at full speed. Instead, individuals are posited to differ in (a) recognition of demands on their intellectual/attentional system, as well as (b) the resultant outcome of a cost–benefit analysis of whether the intellectual/attentional system should be engaged, and if so, by how much. Discrepancies between maximal and typical intelligence might serve as *prima facie* evidence for the plausibility of such underlying processes.

By focusing on the stimulus constellations that provoke high and low intellectual engagement, it may be possible to provide diagnostic information about potential remedies for inappropriate strategies for information processing and suboptimal learning. That is, the existence of further ATIs that affect engagement of intellectual processing may be explored, assessed, and possibly used to provide any necessary remediation.

6. *Developmental consequences.* The literature regarding changes in intellectual ability-as-maximal performance is rife with examples of consistent gains or losses in relative intelligence from childhood into early adulthood (excluding issues of senescence or adult aging-related changes in intellectual performance; e.g., Bayley, 1968). Several developmental theories have been put forth that attempt to explain such consistent changes in relative standing on intelligence tests, not the least of which has to do with environmental privilege. However, some researchers, such as Nicholls and his colleagues (Dweck, 1986; Dweck & Leggett, 1988; Nicholls, 1984; Nicholls, Cheung, Lauer, & Patashnick, 1989; Nolen, 1988), have suggested that individual differences in orientation toward the changeable or constant nature of intelligence are at least partly responsible for gains and losses of relative standing on ability-as-maximal performance measures (see Kanfer, 1990a, for a review). Such concepts have been referred to as implicit *performance orientation* (i.e., ability is believed to be constant, and in order to demonstrate one’s ability, performance should be accomplished with minimal effort expenditures) or *learning orientation* (i.e., ability is changeable, especially as a result of continued effort expenditures in learning tasks). Although there is little long-term developmental evidence for such propositions, the concepts of performance and learning orientations are consistent with the concept that individuals differ, at least in their self-reports, in desired effort expenditures during task engagement. This is a factor that would lead to some learners showing small discrepancies between intelligence-as-maximal and intelligence-as-typical performance (those with

learning orientation) while other learners would have potentially large discrepancies between intelligence as maximal and typical performance (those with performance orientation).

The fundamental proposition here, as with some theories of motivation, is that the individual has some degree of *choice* in how much effort is devoted to intellectual activities. Such choices, when aggregated over substantial time and situation, will have a potentially significant impact on subsequent measures of ability-as-maximal performance, given the relationship between effort expended and the declarative stages of knowledge and skill acquisition.

Beyond Intelligence-as-Typical Performance

Given the importance of “situation” on responses (from the ATI and personality literatures), it could be claimed that no single situation (or set of situations) can be abstracted to provide a single estimate of the multidimensional construct of intelligence. This point seems consistent with Cronbach’s (1990) view of the limiting nature of categorizing any measure into an assessment either as typical or maximal performance. To move beyond the two notions of intelligence-as-maximal and intelligence-as-typical performance, it may be ultimately more promising to identify and sample most critical families of situations, and provide a map of an individual’s performance under the various situations. Identification of these situations, in and of itself, represents a challenging research issue.

Structural Implications of the Proposed Theory

Given the fact that much of the extant theory of intelligence is predicated on the structure of intellectual abilities, it is appropriate to ask how the current notions about attention and intelligence, and the contrast between intelligence-as-maximum performance and intelligence-as-typical performance interact with specific classes of abilities. One implication is that intellectual abilities can be organized along the dimension of attentional demands, from highly resource-dependent to highly resource-insensitive. That is, tests of intellectual abilities that tap highly learned declarative or procedural knowledge will likely involve reduced demands on attentional allocations, and thus be more insensitive to differences between an individual’s maximal and typical attentional/intellectual engagement. Depending on the population of test takers, test that have reduced attentional demands might include those of simple numerical processes, such as addition and subtraction, general information, aspects of perceptual speed ability, psychomotor ability, and so on. Under the rubric proposed here, performance on such ability measures will be less influenced by manipulations

which yield maximal or typical performance.

On the other hand, tests that tap reasoning, working memory, dealing with novel or inconsistent stimuli, will more likely be attentional resource-dependent. As such, large discrepancies in test performance (for at least some examinees) are expected when an exam is administered under maximal or typical intellectual engagement conditions.

Fluid and crystallized intelligence. At a surface level of analysis, there are some sources of overlap between the constructs of resource-insensitive/resource-dependent information processing and the Horn and Cattell (1966) constructs of crystallized/fluid intelligence. Many tests that are primarily resource-dependent would be considered as measures of fluid intelligence. Similarly, many tests identified as being relatively resource-insensitive (such as general information) are considered as measures of crystallized intelligence. However, there are some salient differences between these classificatory schemes, as follows:

According to Horn (1965), "Gc can be measured in tests measuring awareness of concepts, facility and quickness in the use of concept labels and in various reasoning tasks involving cultural concepts and generalized solution instruments" (p. 309). By including reasoning tasks, though they are bound to "cultural" information; for example, the concept of Gc is at odds with the construct of resource-limited tasks. Furthermore, much of the construct of resource-insensitive information processing is not subsumed under Gc at all, but rather under a general speediness factor (Gs). Thus, Gc neither completely subsumes, nor is subsumed within the construct of resource-insensitive information processing.

When tests of Gf are considered, most seem concordant with conceptions of resource dependence. However, the Gf construct does not encompass all types of resource-dependent tests (e.g., the "reasoning involving cultural concepts" listed above, and several aspects of the spatial domain, that Horn & Cattell incorporate into the *Gv* ability). As such, *Gf* can best be thought of as a *subset* of the domain of resource-dependent information processing.

FUTURE DIRECTIONS FOR EMPIRICAL RESEARCH

How Should Intelligence-as-Typical Performance Be Assessed?

Determinants of intelligence-as-typical performance, as contrasted with intelligence-as-maximal performance, are likely to be varied and multidimensional. From the review provided above, several key personal constructs are seen as potential mediators of both current maximal/typical discrepancies, and prospective developmental changes to the differences between typical and maximal

performances. Such constructs might include personality characteristics, such as self-monitoring, performance vs. learning orientation, and achievement motives (e.g., see Kagan, Sontag, Nelson, & Baker, 1958, regarding personality and IQ change).

A second source of influence is probably captured by interest measures, or such measures as are often included in so-called biographical data surveys. For example, interests in cognitively demanding activities (such as reading, debating, problem solving and reasoning games, crossword puzzles, etc.) may very well tap into an individual's preferences for intellectual engagement and thus help describe potential discrepancies between typical and maximal performance levels.

Another major source of influence for typical-maximal discrepancies is likely to be individual differences in self-regulatory skills (e.g., see Kanfer, 1990; Kanfer & Ackerman, 1989). As salient determinants of learning efficacy, individual differences in self-monitoring, self-evaluation, and self-reinforcement processes are likely candidates for influencing a learner's perceptions of the need to engage attentional processes and for maintaining attention over protracted situations when there is little external feedback.

Finally, as Fiske and Butler (1963) recommended *vis-à-vis* personality testing, naturalistic observation may also prove to be beneficial in assessing intelligence-as-typical performance. Discrepancies between typical and maximal intellectual performance seem much more plausible in capturing the "overachievement/underachievement" characterizations that have been attached by achievement test and intelligence differences, given that many existing achievement tests indiscriminantly tap performance on both resource-dependent and resource-insensitive tasks.

Experimental Methods

Further evidence regarding the proposed theory of intelligence could conceivably be derived by experimental methods that affect the individual's intentions and desires for exerting maximal or typical levels of attentional effort. To date, goal-setting studies have shown special promise in narrowing the gap between typical attentional engagement and actual attentional engagement (e.g., Kuhl, 1985; Kuhl & Kraska, 1989). Numerous other ATI paradigms can also be applied to exploring changes to performance resulting from structuring the instructional format or content. Many current computer-assisted instruction systems have built-in capabilities for "engaging" maximal performance. However, too little attention has been given to exploring the consequences of removing or adding these features. Ultimately, through such parametric research, it may be possible to comprehensively outline the types of interventions that attenuate or exacerbate differences between maximal and typical performance, as well as

to delineate which learners are most likely to show large discrepancies between their typical and maximal performance levels, *ceteris paribus*.

Postscript

In the intervening period between the writing of this chapter and its publication, we have created a set of self-report scales of typical intellectual engagement and studied the correlations between these scales and measures of ability and personality (Goff & Ackerman, 1992). As predicted, measures of typical intellectual engagement correlated positively with tests of crystallized intelligence and were uncorrelated with measures of fluid intelligence. In addition, significant overlap was found between measures of typical intellectual engagement and the broad personality factor of openness. These results are supportive of the theory outlined in this chapter and provide an integrated perspective on the interrelations among personality and intelligence constructs.

REFERENCES

- Ackerman, P. L. (1986, April). *Attention, automaticity and individual differences in performance during task practice*. Symposium paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin*, *102*, 3–27.
- 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. (1990). A correlational analysis of skill specificity: Learning, abilities, and individual differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 883–901.
- Ackerman, P. L., & Kyllonen, P. C. (1991). Trainee characteristics. In J. E. Morrison (Ed.), *Training for performance: Principles of applied human learning* (pp. 193–229). West Sussex, England: Wiley.
- Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin*, *101*, 41–74.
- Atkinson, J. W. (1974). Motivational determinants of intellectual performance and cumulative achievement. In J. W. Atkinson & J. O. Raynor (Eds.), *Motivation and achievement* (pp. 389–410). New York: V. W. Winston.
- Bayley, N. (1949). Consistency and variability in the growth of intelligence from birth to eighteen years. *Journal of General Psychology*, *75*, 165–196.
- Bayley, N. (1968). Behavioral correlates of mental growth: Birth to thirty-six years. *American Psychologist*, *23*, 1–17.
- Binet, A., & Simon, T. (1905/1961). The development of intelligence in children. New

- methods for the diagnosis of the intellectual level of subnormals. The development of intelligence in the child. Translated and excerpted in J. J. Jenkins & D. G. Paterson (Eds.), *Studies in individual differences: the search for intelligence* (pp. 81–111). New York: Appleton-Century-Crofts.
- Butler, J. M., & Fiske, D. W. (1955). Theory and techniques of assessment. *Annual Review of Psychology*, 6, 327–356.
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54, 1–22.
- Cronbach, L. J. (1949). *Essentials of psychological testing*. New York: Harper.
- Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). New York: Harper & Row.
- Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.
- Dreger, R. M. (1968). General temperament and personality factors related to intellectual performances. *Journal of Genetic Psychology*, 113, 275–293.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41, 1040–1048.
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256–273.
- Ferguson, G. A. (1954). On learning and human ability. *Canadian Journal of Psychology*, 8, 95–112.
- Ferguson, G. A. (1956). On transfer and the abilities of man. *Canadian Journal of Psychology*, 10, 121–131.
- Fiske, D. W., & Butler, J. M. (1963). The experimental conditions for measuring individual differences. *Educational and Psychological Measurement*, 23, 249–266.
- Fitts, P. M., & Posner, M. I. (1967). *Human performance*. Belmont, CA: Brooks/Cole.
- Flory, C. D. (1940). The intellectual growth of college students. *Journal of Educational Research*, 33, 443–451.
- Goff, M., & Ackerman, P. L. (1992). Personality-intelligence relations: Assessing typical intellectual engagement. *Journal of Educational Psychology*, 84, 537–552.
- Gordon, L. V., & Durea, M. A. (1948). The effect of discouragement on the revised Stanford-Binet scale. *Journal of Genetic Psychology*, 73, 201–207.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Helmreich, R. L., Sawin, L. L., & Carsrud, A. L. (1986). The honeymoon effect in job performance: Temporal increases in the predictive power of achievement motivation. *Journal of Applied Psychology*, 71, 185–188.
- Henry, R. A., & Hulin, C. L. (1987). Stability of skilled performance across time: Some generalizations and limitation on utilities. *Journal of Applied Psychology*, 72, 457–462.
- Honzik, M. P., MacFarlane, J., & Allen, L. (1948). The stability of mental test performance between two and eighteen years. *Journal of Experimental Education*, 17, 309–324.
- Horn, J. L. (1965). *Fluid and crystallized intelligence: A factor analytic study of the structure among primary mental abilities*. Unpublished doctoral dissertation, University of Illinois, Urbana.
- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crys-

- tallized general intelligences. *Journal of Educational Psychology*, 57, 253–270.
- Hulin, C. L., Henry, R. A., & Noon, S. L. (1990). Adding a dimension: Time as a factor in the generalizability of predictive relationships. *Psychological Bulletin*, 107, 328–340.
- Humphreys, L. G. (1979). The construct of general intelligence. *Intelligence*, 3, 105–120.
- Humphreys, L. G., & Taber, T. (1973). Postdiction study of the GRE and eight semesters of college grades. *Journal of Educational Measurement*, 10, 179–184.
- 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.
- Husen, T. (1951). The influence of schooling upon IQ. *Theoria*, 17, 61–88.
- Kagan, J., Sontag, L. W., Nelson, V. L., & Baker, C. T. (1958). Personality and IQ change. *Journal of Abnormal and Social Psychology*, 56, 261–266.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Kanfer, R. (1987). Task-specific motivation: An integrative approach to issues of measurement, mechanisms, processes, and determinants. *Journal of Social and Clinical Psychology*, 5, 251–278.
- Kanfer, R. (1990a). Motivation and individual differences in learning: An integration of developmental, differential and cognitive perspectives. *Learning and Individual Differences*, 2, 221–239.
- Kanfer, R. (1990b). Motivation theory in industrial and organizational psychology. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology, Volume 1: Theory in industrial and organizational psychology* (pp. 75–170). Palo Alto, CA: Consulting Psychologists Press.
- Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology Monograph*, 74, 657–690.
- Kuhl, J. (1981). Motivational and functional helplessness: The moderating effect of action versus state orientation. *Journal of Personality and Social Psychology*, 40, 155–170.
- Kuhl, J. (1985). Volitional mediators of cognition–behavior consistency: Self-regulatory processes and action versus state orientation. In J. Kuhl & J. Beckmann (Eds.), *Action control: From cognition to behavior* (pp. 101–128). New York: Springer-Verlag.
- Kuhl, J., & Kraska, K. (1989). Self-regulation and matamotivation: Computational mechanisms, development, and assessment. In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology: The Minnesota Symposium on learning and individual differences* (pp. 343–369). Hillsdale, NJ: Erlbaum.
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity?! *Intelligence*, 14, 389–433.
- Kyllonen, P. C., & Woltz, D. J. (1989). Role of cognitive factors in the acquisition of cognitive skill. In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology: The Minnesota Symposium on learning and individual differences* (pp. 239–280). Hillsdale, NJ: Erlbaum.
- Lin, P. C., & Humphreys, L. G. (1977). Predictions of academic performance in gradu-

- ate and professional school. *Applied Psychological Measurement*, 1, 249–257.
- Lynn, R., & Gordon, I. E. (1961). The relation of neuroticism and extroversion to intelligence and educational attainment. *British Journal of Educational Psychology*, 31, 194–203.
- Nicholls, J. G. (1984). Achievement motivation: Conception of ability, subjective experience, task choice, and performance. *Psychological Review*, 91, 328–346.
- Nicholls, J. G., Cheung, P. C., Lauer, J., & Patashnick, M. (1989). Individual differences in academic motivation: Perceived ability, goals, beliefs, and values. *Learning & Individual Differences*, 1, 63–84.
- Nolen, S. B. (1988). Reasons for studying: Motivational orientations and study strategies. *Cognition and Instruction*, 5, 269–287.
- Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, 7, 44–64.
- Parasuraman, R. (1984). Sustained attention in detection and discrimination. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 243–271). Orlando, FL: Academic Press.
- Quereshi, M. Y. (1960). Mental test performance as a result of payoff conditions, item difficulty, and degree of speeding. *Journal of Applied Psychology*, 44, 65–77.
- Reason, J. (1990). *Human error*. New York: Cambridge University Press.
- Reason, J., & Mycielska, K. (1982). *Absent-minded? The psychology of mental lapses and everyday errors*. Englewood Cliffs, NJ: Prentice-Hall.
- Revelle, W. (1986). Motivation and efficiency of cognitive performance. In D. R. Brown & J. Veroff (Eds.), *Frontiers of motivational psychology: Essays in honor of John W. Atkinson* (pp. 107–131). Berlin: Springer-Verlag.
- Revelle, W. (1989). Personality, motivation, and cognitive performance. In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology: The Minnesota Symposium on learning and individual differences* (pp. 297–341). Hillsdale, NJ: Erlbaum.
- 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.
- Sattler, J. M., & Theye, F. (1967). Procedural, situational, and interpersonal variables in individual intelligence testing. *Psychological Bulletin*, 68, 347–360.
- Snow, R. E. (1989a). 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. 13–59). New York: W. H. Freeman.
- Snow, R. E. (1989b). Cognitive-conative aptitude interactions in learning. In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology: The Minnesota Symposium on learning and individual differences* (pp. 435–474). Hillsdale, NJ: Erlbaum.
- Snow, R. E., & Yalow, E. (1982). Education and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 493–585). Cambridge, UK: Cambridge University Press.
- Spearman, C. (1927). *The nature of "intelligence" and the principles of cognition*. New York: Macmillan.

- 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 acquisition. In P. L. Ackerman, R. J. Sternberg, & R. Glaser (Eds.), *Learning and individual differences: Advances in theory and research* (pp. 255–278). New York: W. H. Freeman.
- Thorndike, E. L., Bregman, E. O., Cobb, M. V., Woodyard, E., & the staff of the Division of Psychology of the Institute of Educational Research of Teacher's College, Columbia University. (1926). *The measurement of intelligence*. New York: Teachers College, Columbia University.
- Thorndike, R. L. (1940). "Constancy" of the IQ. *Psychological Bulletin, 37*, 167–186.
- Vernon, P. E. (1950). *The structure of human abilities*. New York: Wiley.
- Wechsler, D. (1944). *The measurement of adult intelligence*. Baltimore: The Williams & Wilkins Co.
- Whitehead, A. N. (1911). *Introduction to mathematics* (American ed., revised and reset, 1948). New York: Oxford University Press.
- Wickens, C. D. (1980). The structure of attentional resources. In R. S. Nickerson (Ed.), *Attention and performance* (Vol. VIII, pp. 239–257). Hillsdale, NJ: Erlbaum.
- Woltz, D. J. (1988). An investigation of the role of working memory in procedural skill acquisition. *Journal of Experimental Psychology: General, 117*, 319–331.
- Woodrow, H. (1946). The ability to learn. *Psychological Review, 53*, 147–158.
- Zeaman, D. (1978). Some relations of general intelligence and selective attention. *Intelligence, 2*, 55–73.
- Zeaman, D., & Kaufman, H. (1955). Individual differences and theory in a motor learning task. *Psychological Monographs: General and Applied, 69*(6, Whole No. 391).

Chapter 2

***Intelligence, Personality,
and Society: Constructivist
versus Essentialist Possibilities***

Chris Brand

Vincent Egan

Ian Deary

*Department of Psychology
University of Edinburgh*

Scotland is a country that is better known for its engineering than for its philosophical genius. It may seem strange to say this. Scotland gave birth to David Hume, whose skeptical talent was long to prove a match for the associationist realism of England's John Locke and the dissociationist idealism of Ireland's Bishop Berkeley. Yet Scotland did little to honor the atheistic, Tory devotee of English ways that Hume was to become. It refused Hume a chair in any of its four universities and let him pass his days as Keeper of the Advocate's Library in Edinburgh.

Hume, of course, fell out with his fellow Scots about the faculties of the mind and about the possibilities of reasonably secure human knowledge,

notably about causation and about the real, enduring existence of persons. In 18th-century Scotland's Enlightenment there was a belief that the essential faculties of the mind (as they were first conceived by Aristotle) were able (with some help from holy scripture in normal circumstances) to deliver truth. In particular, Scottish philosophers, drawing on St Augustine, developed the notion that we all have a special faculty of conscience that is able to detect moral truth. Hence, conveniently for the aspirations of Scotland's (Reformed) Kirk, it was thought possible to find some true and just way of ordering society (MacIntyre, 1988) in the corporatist style that the Scots were long to prefer.

Hume, by contrast, could see none of these possibilities. Enamored of English society that had settled for a middle way of compromise and passable pluralism (within the Church of England) in matters of religion and for dispersed property ownership as a way of life, Hume doubted that strict Protestantism and the disciplined society that it entailed could be vindicated, at least to the satisfaction of the questioning philosopher. Doubting even the existence of his own self (he could see inside him no such entity, only "bundles of sensations"), it was Hume who would awake Königsberg's Kant "from his dogmatic slumbers." Hume was to provide the demolition job from which Hegel would finally build a strictly German (less Kantianly "transcendental") idealism. Hegel, in turn, would begin the long tradition of subjectivism in philosophy and social thought that would reach through Schopenhauer, Nietzsche, Heidegger, and Sartre down to the Frankfurt School, Jürgen Habermas, and the present day.

Presently, the notion that the important aspects of human life and human psychology consist chiefly in ideas finds its more familiar expression in the constructivism that has proved popular with some developmental and social psychologists (Bruner, 1986; Gergen & Shotter, 1989) and in the denials of "person-as-essence" of some philosophers (Parfit, 1984). Such idealism (or sometimes nominalism, envisaging person terms as a mere "language game") about the person and about personality normally seems to have special attraction for would-be social and political reformers for a reason that can easily be identified. If there are real people with essential, enduring properties who live in a real world that science can but painstakingly discover, what is the task of the social reformer and champion of the underdog? Why, it will be to identify (and even measure) people's properties and psychopathologies and trace them to their often quite distant genetic and/or environmental sources. Now such an undertaking requires discipline, persistence, and perhaps even the quantitative method. In any case, the task has long been attempted by 20th-century psychologists with outcomes that have proved disagreeable for the reformer. Many convenient environmentalistic ideas (that social class is strongly causal to children's intelligence, that maternal deprivation causes unhappiness and life failure) have been dashed from the hands of enthusiasts (e.g., Brand 1987a; Scharfstein, 1980). The major environmentalistic effects that can be

detected operate on such a time scale as to offer faint hope of intervention. (No one currently knows why IQ-type levels on quite a few mental tests have been rising through the 20th century [Brand, 1987b].) Worse, the alternative prospect that genetic factors account importantly for human psychological realities appears deeply divisive and pessimistic (at least until IQ-boosting drugs or genetic engineering methods are discovered). How much more agreeable it would be if people consisted chiefly in their ideas, perceptions, constructions, language games, and so forth. For all these features seem much more open to overnight “deconstruction” and “reconstruction” just as the revolutionaries of 1789 and 1917 once hoped.

Whatever the precise advantages of idealism, and whatever the dangers (Bloom, 1987) that unconstrained idealism may lead to a new yearning for a reconstructing Superman with a new social mythology (as it did once in its German heartland), it must be admitted that an alternative, essentialist psychology has yet to find its feet. Hans Eysenck—for many years the venerated leader of the London School of psychology—has long advocated that many major human individual differences in personality and attitudes be traced primarily to a small number of underlying differences, notably to dimensions of difference that he has named Psychoticism, Neuroticism, and Extraversion. The labor expended in this 40-year exercise has been immense, and could profitably have been still greater. Yet, today, there is still, despite whatever hopes, no widely agreed scientific story that can be told of P, N, or even E (on which most work has been done: Brand, 1983; Eysenck, 1981; Eysenck & Eysenck, 1985; Neiss, 1990; Zinbarg & Revelle, 1989). Even in psychogenetic studies, where passably certain methods exist for tracing human differences to genetic and/or environmental origins, the story that emerges of Eysenck’s dimensions shows many peculiarities that must displease the traditional champion of nomothetic approaches to personality. Though identical twins do appear to be substantially alike on Eysenck’s dimensions, there are few similarities in personality between other types of relatives even when genetic overlap is 50% and when the individuals have grown up together (Brand, 1989; Eaves, Eysenck, & Martin, 1989; see Table 1).

By contrast, there is one area of the London School’s enquiries that has yielded a clear nomothetic success story. This is in the study of intelligence, where work of the 1980s continued to confirm that general intelligence (*g*) is markedly unitary, measurable with fairness, importantly predictive, based to some extent in elementary cognitive processes (such as inspection time), underpinned by detectable psychophysiological processes, substantially heritable by normal, additive genetic mechanisms, yet influenced too by social environment to a detectable extent (Bouchard, Lykken, McGue, Segal, & Tellegen; Brand, 1987a, 1990, 1992; Brand, Caryl, Deary, Egan, & Pagliari, 1991). (Just about the only major scientific problem with IQ at present is that of explaining its apparent tendency to secular increase, as Flynn [1987] has been eager

Table 1. Correlations Between Relatives for Personality Traits

	<i>Psychoticism</i>	<i>Extraversion</i>	<i>Neuroticism</i>	<i>Lie</i>
MZ _m twins (N = 70 pairs)	0.53	0.65	0.51	0.53
MZ _f twins (N = 233 pairs)	0.41	0.46	0.45	0.56
DZ _m twins (N = 47 pairs)	0.16	0.25	0.03	0.02
DZ _f twins (N = 125 pairs)	0.38	0.18	0.09	0.52
Brothers (N = 72)	0.08	0.16	-0.01	0.26
Sisters (N = 151)	0.22	0.36	0.07	0.38

Extracted from *Genes, Culture and Personality: An Empirical Approach*, by L. J. Eaves, H. J. Eysenck, & N. G. Martin, 1989, London: Academic Press. Reprinted by permission.

to point out. For an explanation in terms of speed-oriented test-taking strategies, see Brand, 1987b, 1990b.)

In view of the widely diverging rates of success for, on the one hand, the London School's approach to personality and, on the other hand, its approach to intelligence, it is perhaps worth asking whether intelligence (which alone allows much of distinctively human learning, contrary to the dreams of generations of experimental and now cognitive psychologists) could not be made to do rather more work than is conventionally expected by psychometricians, always prone to defer to their experimentalist masters. Perhaps intelligence can provide explanations and clarification of personality structure when other putative mechanisms have failed.

There are four supportive considerations:

1. It was Socrates who first identified what may be called "the unity of the virtues"—the commonplace phenomenon that it is hard to imagine possessing one characterological virtue (honesty, courage, kindness, carefulness) without also possessing some of the others to some extent. Subsequently Aristotle's account of personality came to focus on the possibility that reason, with the help of education, could come in the course of personality development to foster virtue and master the passions. Later, Christian moral theology invariably took the battle for reason to achieve supremacy over the passions as a major theme. The godless Freud saw matters similarly. In the 20th century, the frequency with which factors of *emotionality* (Brand, 1984) and, sometimes, *sense versus sensibility* (Brand & Egan, 1989; Table 2) appear in self-report indicators of personality continue to suggest that people's experienced balance of reason and sense versus mood and emotion is a major matter, of which they are keenly aware and can report with considerable reliability. Whether or not

Table 2. Adjectives Defining One of the Largest (post-scrree-test) Rotated Factors in a Study of Ipsative, Adjectival Self-descriptions by Edinburgh Psychology Students

<i>Sense</i>	<i>Sensibility</i>
cool-headed	loving
thick-skinned	passionate
level-headed	compassionate
composed	wilful
deliberate	hot-blooded
unsentimental	tender-hearted
	excitable
	hasty
	soft-hearted

From "The 'Big Five' Dimensions of Personality? Evidence from Ipsative, Adjectival Self-Attributions," by C. R. Brand & V. Egan, 1989, *Personality and Individual Differences*, 10, 1165–1171. Reprinted by permission.

they can assess their own intelligence, they have a clear idea of whether their intelligence manages efficiently to harness their passions.

2. That intelligence is viewed as a necessary condition of important human qualities is readily seen in our discourse. Of Alzheimer's disease, one writer claims: "As the disease progresses, all aspects of intelligence are affected and the individual's personality is lost." Some people, fearing such possibilities, wish to be able to make advance directives for their own euthanasia in the event of grave intellectual impairment, whereas currently, people do not seem to worry so much about the listlessness, conservatism of opinionation, and attachment to socialized medicine that are also not infrequently associated with aging.

3. Our courts of law in the West also acknowledge the importance of intelligence to underpinning capacities for the exercise of free will and moral responsibility: A low mental age is the most commonly used rationale for deciding that a person did not really "intend" a serious crime—children are deemed not to have *mens rea* until around age 11. Of course, such judgments of mental age are sometimes made indirectly and conveniently via the "hard" public variable of chronological age. However, recently, the Supreme Court of Victoria, Australia, used evidence from mental testing to decide that a grossly cerebrally palsied teenage girl, who had hardly any voluntary motor control, did indeed have a "will" of her own and could thus be considered able to decide her own future (outside of, as she wanted it to be, the squalid hospital for the grossly handicapped and supposedly mentally retarded children to which she had been consigned).

4. In the layperson's view of intelligence, many valued human qualities (e.g., showing interest in and affiliation with one's peers, and exhibiting inde-

Table 3. Five Groups of Personality Features Found to Be Positively Related to Intelligence in Empirical Work (see Brand, 1984, for references)

<i>Features</i>	<i>Typical correlation with g</i>	<i>Brand's nomenclature</i>
anxiety, moodiness, low self-esteem	-0.30	neuroticism (<i>n</i>)
fluency, creativity, expressiveness	+0.40	energy (<i>e</i>)
moral development, restraint, conscientiousness	+0.30	conscientiousness (<i>c</i>)
independence, initiative, assertion	+0.40	will (<i>w</i>)
tender-mindedness, benevolence, social interest	+0.30	affection (<i>a</i>)

pendence, creativity and open-mindedness) are held to be substantially associated with it (Sternberg, 1990).

So there is some reason to believe that intelligence is important to nonpsychologists when they consider personality. Nevertheless, can intelligence as psychologists measure it be related demonstrably to important personal powers? Binet certainly thought so: He said intelligence found its outer boundaries in (was "constrained by") "invention, direction, criticism and comprehension"—or, as we might say today, in creativity, initiative, conscientiousness, and open-mindedness. But what does empirical study show? One of us (Brand, 1984) reviewed some fraction of the relevant literature a few years ago and found, for each of Binet's propositions, broadly understood, some half-dozen studies attesting to them, with correlations of around +0.35 (Table 3). For example, measures of "creativity" administered to subjects of a normal range of IQs correlate at around +0.40. More widely, going beyond such general personality features, a second survey of the literature (Brand, 1987d) found IQ to be associated with some 40 human characteristics that are often naively presumed to be nonintellective (Table 4). Thus IQ is associated with altruism rather than with psychoticism, with leadership rather than with acquiescence, with moral reasoning as opposed to delinquency, and with suffering anorexia nervosa rather than hysteria.

Again, though these matters are insufficiently investigated, it is readily acknowledged that IQ will play some part in musical preferences, dietary taste, and dress sense. At least some types of IQ estimate are known to be associated with holding broadly liberal social values: This was recently confirmed by one of us in Edinburgh, with a correlation of +0.45 from a normal sample of young people, including a considerable proportion on the borderline of unemployment (Egan, 1989); and IQ is the only substantial psychological predictor of marital choice. Such is the relevance of IQ to socially approved conduct that one American sociologist (Gordon, 1987) has recently argued in some detail

Table 4. Examples of Personal Features Found to Be Positively Related to Intelligence (see Brand, 1987d, for references)

Achievement motivation
Altruism
Art preferences
Creativity
Emotional awareness
Field independence
Humor, sense of
Interests, breadth and depth of
Involvement in school activities
Leadership
Moral reasoning and development
Musical preferences and abilities
Psychotherapy, response to
Sports participation at university
Supermarket shopping ability
Talking speed
Values (especially 'liberalism')

that the overinvolvement of American blacks in crime can be entirely accounted for by their measured level of IQ on conventional tests. Possibly, then, psychologists of personality have ignored for too long evidence of IQ's involvement in personality that was available to them if they had cared to look for it. They have preferred to search haystacks for evidence of criminality being traceable to Pavlovian unconditionability, to any of Piaget's or Kohlberg's "stages," to cognitive-psychological black boxes that have so far done nothing to illumine personality, and to sociological variables that have proved quite unable to explain the very modest rise in crime that accompanied Britain's massively increased levels of unemployment after 1977. Arguably, *g* is to the psychology of personality what carbon is to chemistry: *g* accounts for more variance in important human behaviors than do all the other variables in anthropology, sociology, and experimental psychology put together.

Why, then, have psychologists been reluctant to entertain this? Undoubtedly the biggest problem here is that people themselves are far from being realists in their own direct estimations of their own intellectual levels. When, in Edinburgh, we asked 47 psychology students to rate themselves on adjectives such as *quick*, *clever*, *bright*, *knowledgeable*, and *skillful* we found no correlation at all between self-rated intelligence and their IQ test performance. Instead, we found a strong negative correlation (-0.60) between self-rated intelligence and other self-ratings of affection and tender-mindedness. (The ipsative adjective checklist that was used is reported by Brand & Egan, 1989.) American researchers report broadly similar findings (Harter, 1990): Self-ratings of competence bear little relation to IQ but, instead, negative relations to self-assessed emotionality and anxiety. Clearly, if people of high IQ are often too modest to admit their own intelligence, the construct is not likely to appeal

to them as an explanation of their moral development or romantic choices.

Why is g 's influence thus unrecognized by such testees? There would seem to be four possible ways of explaining this strange phenomenon:

1. It may be that people—and especially students—today are so much in thrall to egalitarian dogma (denying the importance of intellectual differences) that they politely decline to admit their own abilities. Conversely, some people, at lower levels of IQ, may accept inflated estimates of their own ability levels and thus draw illusory consolation from egalitarian ideology.

2. In the modern West, people live their lives with quite a limited range of friends and associates of around their own intellectual and educational levels. Gone, at least in 20th-century Britain, are the days of students mixing with a wide range of humanity at their churches or local football clubs. Perhaps, in such circumstances of milieu choice along the lines of IQ, intelligence becomes an irrelevant variable to much of ordinary life: It will no longer provide much by way of local explanation of why one differs from one's own self-chosen peers.

3. A more philosophical possibility is simply that it is very difficult for people to admit their own limitations in intelligence—unless they do so out of a wider, more generally anxious disposition. Perhaps, no more than Descartes, can one rationally doubt one's own capacity to think—for, if one did, one would by that same token not even be able to judge soberly that one was to any serious degree deficient in intelligence.

4. A more interesting possibility, however, is that at higher levels of intelligence, personality itself develops and differentiates in ways that make g recede somewhat into the background as a serious determinant of what the life of the person is about. Perhaps, at higher g levels, people are to be found interested and invested in the particular values, motivations, interests, and sensibilities that g has so distinctly developed for them, rather than in the dull fact that, without g , they would not have their distinguishing proclivities and specialties at all. Thus, for example, it is notoriously difficult to measure, with any serious reliability, the personalities of children: Even by age 8, only neuroticism and extraversion are tolerably measurable (with whatever considerable allowance for massive 'faking good' by many children), and this is a good three years after IQ itself has already assumed its own lasting predictive potency. Again, notoriously, the number of reliable personality dimensions that can be retrieved from amongst students in tertiary education is higher than is found in younger, less intelligent children or in normal samples of the adult population. (Eysenck's two dimensions of N and E long seemed adequate to describing personality variation in normal population samples, while American investigators, dealing largely with more educated subjects, found more dimensions to be necessary.)

In Edinburgh, we have recently tried out this possibility in a limited way. Administering our adjectival self-rating device to 35 lower and 27 higher IQ

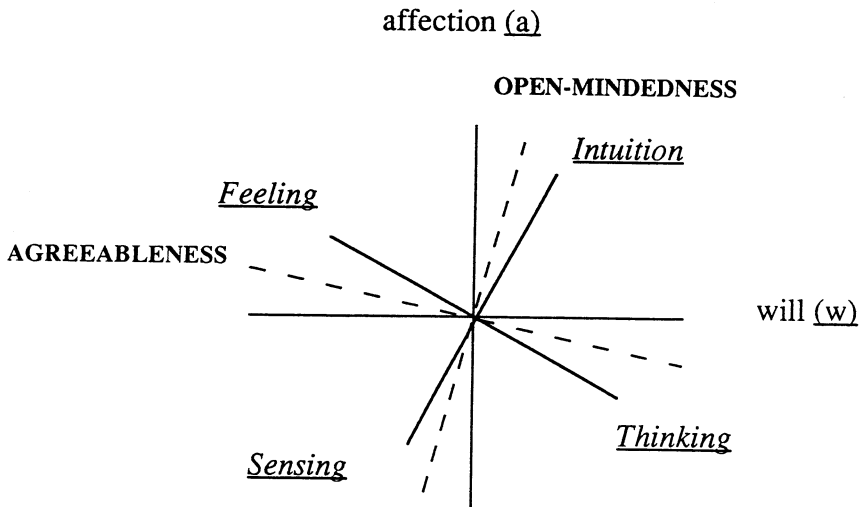


Figure 1. Suggested relations between two dimensions that appear in the "Big Six," the "Big Five" and in the Myers-Briggs Type Indicator.

students, we asked whether more dimensions were necessary to account for variance in the latter group. For example, we asked whether one broad dimension of "extraversion (*e*) versus conscientiousness (*c*)" (realizing Eysenck's historic aspiration for a dimension that would oppose extraversion to conventional morality) would prove adequate to differentiating among the lower IQ students. By contrast, we asked whether two relatively independent dimensions of extraversion *and* conscientiousness would prove necessary to describe variation amongst higher IQ students.

In fact, no such distinction was necessary, rather the reverse. It was our higher IQ group that showed a strong negative correlation (-0.72) between *e* and *c*, although our lower IQ students showed no significant correlation ($r = +0.08$). Evidently it was our higher IQ students who provided the best instantiation of the original Eysenckian vision: They had definitely polarized as to whether they were extraverted *or* conventionally conscientious. In our sample (where of course even the "lower" IQ group had a mean IQ of 114) we thus found no evidence for more *dimensions* amongst the higher IQ, but some evidence of more dramatic distinctions, as if the higher IQ students had more sharply defined personalities.

It is regrettably difficult to explore this question further, for much of the copious data of 20th-century psychometry barely allow the question to be posed. However, we have discovered one test instrument on which relevant data have been provided in some quantity. This is the Myers-Briggs Type Indicator

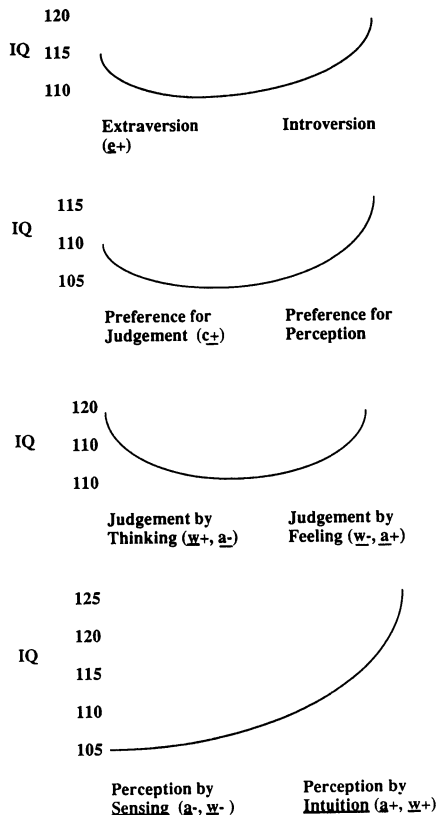


Figure 2. Relations between IQ and the four “Jungian” dimensions of the Myers-Briggs Type Indicator ($N = 5,025$). (Approximate “Big Six” equivalents to these dimensions are indicated, as in Figure 1.)

(Myers, 1962): a questionnaire that purports to assess four Jungian dimensions of personality—dimensions that are known from McCrae and Costa’s (1988) work to be closely related to four of the Big Five (or Six, including g) dimensions of modern psychometric psychology, as is indicated in Figure 1. These dimensions are: Judging (liking routine) versus Perceiving (not minding uncertainty); Extraversion (liking social interaction) versus Introversion; Thinking (liking rational argument) versus Feeling; and Sensing (liking facts) versus Intuition (liking ideas). For most of these dimensions, from data on 5,025 American high school testees, it turns out that there is a U-shaped function (Myers, 1962, p. 92 ff): Higher IQ testees are more likely to be found at the extremes of these dimensions of cognitive style, as is shown in Figure 2.

The six-dimensional model of personality

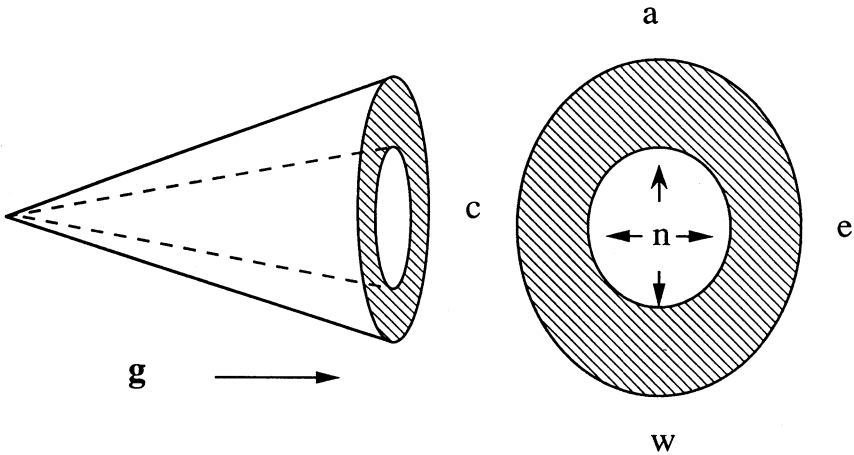


Figure 3. Brand's *mandala*—a double cone (cf. Yeats, 1939) representing how *g* may play out into wider extremes of personality development (with neuroticism (*n*) providing parallel differentiation and variability). Key: *w* = will, *a* = affection, *c* = conscientiousness, *e* = energy, *n* = neuroticism, *g* = general intelligence.

Once again, but with far greater empirical authority, we can say that higher IQ people seem, after a fashion, to have more personality: They are more distinguished one from another, on questions about their own personal styles. This finding is, of course, in line with Jungian expectations that quite marked bipolar choices of style have to be made in the course of adequate personality development.

To truly embrace such findings requires more consideration than is possible in this chapter. One way of thinking about each human personality is that it extends, so long as *g* lasts, over a conic space within what we may call, trying to follow the Irish poet W. B. Yeats (1939), a “double cone” (which we illustrate in Figure 3). So long as *g* is high, we see conspicuous diversity, readily trapped in suitable questionnaires. As has been claimed for intelligence itself (Brand, 1987c; Detterman & Daniel, 1989), this may involve more dimensions (probably depending on how we set about measuring personality); but, certainly in some of our own data and that from the large Myers-Briggs normative sample, it seems to involve higher IQ subjects showing more distinction amongst themselves. This dependence of personality upon intelligence is

allowed in the “double cone” within which higher IQ subjects are predicted to show wider variation in personality. More prosaically, however, we would set personality psychologists to exploring systematically the possibilities that we have indicated: Fresh data sets will be required that involve *g* alongside conventional personality variables.

The notion of the person having “essence” is disputed by psychologists and philosophers who have given up the struggle of trying to discern the ways of the human mind, heart, soul, and spirit, probably because such theorists fail to realize the understanding of these wonders that only analysis that takes the role of *g* seriously will make possible. General intelligence makes not only for personality variation, it also structures that variation in novel ways, allowing new breadths and adumbrations of personality. Intelligence fuels and sustains personality until old age sets in, leaving us all with the merest crystallized remnants that *g*'s fluid power once made seem the height of style. Scottish faculty psychology may have been scorned by a skeptical, renegade Hume; it was subsequently neglected by the Continental idealists who sprang up in its wake; and the Scottish tradition itself had to be reproved by Sir Cyril Burt (on a visit to Edinburgh in 1927) for failing to take *g* seriously enough. But Aristotle's sober, essentialist vision of the biological underpinnings and sustenance of personhood served the west well for 2,300 years and should now be reconsidered with the help that today's firm understanding of *g* and its importance can arguably provide. Modern societies have seen their own mind-analyzing experts fruitlessly at work in the 20th century—recently, despite multibillion dollar expenditures, finding “very little” from presumptive studies of “artificial intelligence” (Aleksander, 1990; Gregory, 1993). It is time to acknowledge the centrality of naturally occurring intelligence to human personality, to human achievement, and to how societies can endeavor to justify and improve themselves.

REFERENCES

- Aleksander, I. (1990). Machine dilemmas. *Nature*, *348*, 399.
- Bloom, A. (1987). *The closing of the American mind*. New York: Simon & Schuster.
- Bouchard, T. J., 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.
- Brand, C. R. (1983). Review of H. J. Eysenck (Ed.), *A model for personality*. *Behaviour Research and Therapy*, *31*(1), 129–131.
- Brand, C. R. (1984). Personality dimensions: An overview of modern trait psychology. In J. Nicholson & H. Beloff (Eds.), *Psychology Survey 5*. Leicester, UK: British Psychological Society.
- Brand, C. R. (1987a). A touch of (social) class. *Nature*, *325*, 767–768.

- Brand, C. R. (1987b). Bryter and still bryter. *Nature*, 329, 110.
- Brand, C. R. (1987c). Review of R. J. Sternberg & D. K. Detterman (Eds.), *What is intelligence? Journal of Social and Biological Structures*, 11, 396–398.
- Brand, C. R. (1987d). The importance of general intelligence. In S. Modgil & C. Modgil (Eds.), *Arthur Jensen: Consensus and controversy*. Brighton, UK: Falmer.
- Brand, C. R. (1989). A fitting endeavour. *Nature*, 341, 29.
- Brand, C. R. (1990b). A 'gross' underestimate of a 'Massive IQ rise'? A rejoinder to Flynn. *Irish Journal of Psychology* 11(1), 52–56.
- Brand, C. R. (1990a). The psychology of psychometric intelligence: agreeable sentiments and awkward correlations. *Personality and Individual Differences*, 11, 535–537.
- Brand, C. R. (1992). Review of R. J. Sternberg & J. Kolligian (Eds.), *Competence Considered. Behaviour Research & Theory* 30(4), 422–423.
- Brand, C. R., Caryl, P. G., Deary, I. J., Egan, J., & Pagliari, H. C. (1991). Is intelligence illusory? *The Lancet*, 337, 428.
- Brand, C. R., & Egan, V. (1989). The 'Big Five' dimensions of personality? Evidence from ipsative, adjectival self-attributions. *Personality and Individual Differences*, 10, 1165–1171.
- Bruner, J. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.
- Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest for low-IQ groups. *Intelligence*, 13, 349–359.
- Eaves, L. J., Eysenck, H. J., & Martin, N. G. (1989). *Genes, culture and personality: An empirical approach*. London: Academic.
- Egan, V. (1989). Links between personality, ability and attitudes in a low-IQ sample. *Personality and Individual Differences*, 10, 997–1001.
- Eysenck, H. J. (1981). *A model for personality*. New York: Springer Verlag.
- Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and individual differences: A natural science approach*. New York: Plenum.
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 171–191.
- Gergen, K. J., & Shotter, J. (Eds.). (1989). *Texts of identity*. London: Sage.
- Gordon, R. A. (1987). SES versus IQ in the race-IQ-delinquency model. *International Journal of Sociology and Social Policy*, 7, 30–95.
- Gregory, R. L. (1993). *Desert Island disks*. London: B.B.C. IV (U.K.).
- Harter, S. (1990). Causes, correlates, and the functional role of global self-worth: a life-span perspective. In R. J. Sternberg & J. Olligian Jr., *Competence considered*. New Haven, CT: Yale University Press.
- MacIntyre, A. (1988). *Whose justice? Which rationality?* London: Duckworth.
- McCrae, R. R., & Costa, P. T., Jr. (1988). The structure of interpersonal traits: Wiggins's circumplex and the five-factor model. *Journal of Personality and Social Psychology*, 56, 586–595.
- Myers, I. B. (1962). *Manual for the Myers-Briggs type indicator*. Palo Alto, CA: Consulting Psychologists Press.

- Neiss, R. (1990). Ending arousal's reign of error: A reply to Anderson. *Psychological Bulletin*, 107, 101–105.
- Parfit, D. (1984). *Reasons and persons*. Oxford, UK: Oxford University Press.
- Scharfstein, B. A. (1980). *The philosophers*. Oxford, UK: Blackwell.
- Sternberg, R. (1990). Prototypes of competence and incompetence. In R. Sternberg & J. Kolligian (Eds.), *Competence considered*. New Haven, CT: Yale University Press.
- Yeats, W. B. (1939). *A vision*. London: Macmillan.
- Zinbarg, R. Z., & Revelle, W. (1989). Personality and conditioning: A test of four models. *Journal of Personality and Social Psychology*, 57, 301–314.

Chapter 3

***Cognitive Abilities:
Constructing a Theory
from Data****

John B. Carroll

University of North Carolina at Chapel Hill

According to most accounts, a theory is a set of statements or hypotheses concerning the explanation of a set of data. In the present case, the data for which a theory is to be constructed are voluminous. They consist, essentially, of information on the (usually Pearsonian) correlations of variables coming primarily from psychological tests, but occasionally (*very* occasionally) from observations or ratings of behavior from sources other than tests, such as school marks, peer ratings of leadership qualities, and the like. More specifically, they come from several hundred studies that I have assembled (Carroll,

* This material is based upon work supported by the National Science Foundation under Grant No. BNS 82-12486. I wish to acknowledge, with thanks, the devoted and highly capable efforts of Christina M. Gullion and Ann C. Meade, research associates during the years 1983 to 1986, who helped in the reanalysis of datasets and in other project tasks. I am also indebted to the numerous investigators whose published and unpublished factor-analytic studies I have reanalyzed.

1993) from the psychological literature—studies that were designed to investigate, usually by means of factor analysis, the dimensions of individual differences in cognitive abilities.

It is almost banal to note that theory tends to guide the collection of data. The selection of what data to collect—what variables to use, what samples to employ, and so on—depends on hypotheses concerning what the data might be expected to reveal. It may be assumed that those who collected the data had various notions as to what kinds of cognitive abilities may be presumed to exist and to be distinguished one from another. These notions—some of them, possibly, derived from common observation and armchair speculation, but others based on a tradition of empirical research in the field of cognitive abilities—were influential in the selection of variables.

THEORY BEHIND THE METHODOLOGY

There was also theory behind the methodology. First consider the methodology of testing, which is based on a theory of abilities, or “latent traits.” The notion is that a psychological test provides an indication of the degree of ability that the individual can demonstrate in performing some class of tasks. It is further assumed that over the population of individuals, or at least in a representative sample of individuals, persons differ in their degrees of ability, and this assumption is apparently confirmed when test score distributions show wide variance. Mental test theory, whether of the “classical” or the “item response theory” kind, assumes that some part of a test score is due to a “true ability,” and provides methods for making the estimation of true ability as accurate as possible. I should also add that the equations of item response theory assume that when the items of a test are arranged in order of difficulty, an individual’s probability of success decreases as difficulty increases. The location of the individual response curve is a function of the individual’s standing on the true ability scale. I have made this fact the basis for focusing on what I call the *person characteristic function*, the function relating probability of passing items to task difficulty and to true ability. I consider that this function is valuable in defining abilities as a function of the characteristics of items or tasks (Carroll, 1987, 1990; Carroll, Meade, & Johnson, 1991).

Second, consider the methodology of factor analysis, which introduces theoretical considerations of its own. Put in the simplest terms, factor analysis assumes that the “true abilities” underlying the variables in a factor-analytic study are or can be different. One object of a factor-analytic study is to determine how many different “true abilities” or latent traits are represented in the variables selected for study. In order to do this, factor analysis assumes a mathematical model that specifies the composition of any given variable as possibly a function of a number of different “factors” or sources of variance. There are

three possible kinds of source variance assumed: common factor variance, specific variance, and error variance. For most purposes, specific and error factors are of little interest in factor analysis. What is of most interest is the number and nature of the common factors, each of which is conceived to be a latent trait of ability. The common factor-analytic model can be expressed in the form of an equation, according to which any variable y is taken to be a linear composite of a series of latent factors:

$$y_{ji}x_{1i} + w_{j2}x_{2i} + \dots + w_{jp}x_{pi} \dots \\ + w_{jm}x_{mi} + w_{js}x_{si} + w_{je}x_{ei}$$

In this equation, y_{ji} is the standard score of individual i on variable j ; x_{pi} is the standard score of individual i on factor p , and w_{jp} is the weight of factor p in producing the scores y_{ji} on variable j . In a well-designed and well-analyzed study, and after rotation of axes to a simple structure, factor weights tend to be either near zero, or strikingly positive (e.g., $>.30$).

Many factor-analytic studies suggest that common factors can vary in the degree of their generality over the variables in a study. The generality of a factor can be represented by the *order* at which it emerges in the analysis. Factors that have strikingly positive weights on only a few variables tend to emerge at the first order; those that have strikingly positive weights for clusters of variables that show some differentiation at the first order tend to emerge at the second order, whereas factors that have strikingly positive weights for all the variables of a study, or for a major grouping of them, can emerge at the third order of analysis. As an example of a study showing factors at three orders of analysis, see Table 1, which presents an orthogonalized factor matrix for the nine subtests of the Holzinger-Crowder Uni-Factor Test Battery (reanalyzed from data assembled by Schutz, 1958). As is seen, all tests have strikingly positive loadings on a third-order general factor; different groups of tests have loadings on each of two second-order factors; and still smaller groups of tests have strikingly positive loadings on one or another of four first-order factors.

Such an analysis is based on a model in which the fundamental equation of factor analysis is expanded to something like the following:

$$y_{ji} = w_{jg}x_{gi} \quad \text{(general factor)} \\ + w_{jp}x_{pi} + w_{jq}x_{qi} \\ \quad \text{(second-order factors)} \\ + w_{ja}x_{ai} + w_{jb}x_{bi} + w_{jc}x_{ci} + w_{jd}x_{di} \\ \quad \text{(primary factors)} \\ + w_{ju}x_{ui} \quad \text{(unique factor)}$$

Here, for purposes of illustration it is assumed that there is one third-order factor (g), two second-order factors (p and q), and four first-order factors (a , b , c ,

**Table 1. Hierarchical Factor Matrix for Dataset SCHU11
(Reanalysis of data assembled by Schutz, 1958)**

Variable	Hierarchical Factor							h^2
	1	2	3	4	5	6	7	
Factor 1: 03:F1 General Intelligence (G); order 3								
Factor 2: 02:F1 Crystallized Intelligence (Gc); order 2								
Factor 3: 01:F1 Verbal Ability (V); order 1								
1 Word Meaning.....	.57	.43	.53	-.02	-.01	.01	-.01	.79
2 Odd Words62	.44	.51	.03	.02	.00	.01	.84
Factor 4: 01:F2 Numerical Facility (N); order 1								
6 Remainders.....	.53	.22	-.01	.64	.18	.04	-.01	.77
5 Mixed Arithmetic.....	.56	.25	.01	.62	.16	-.03	.01	.79
Factor 5: 02:F2 Fluid Intelligence (Gf); order 2								
Factor 6: 01:F3 Spatial Relations (SR); order 1								
4 Hatchets50	.01	.01	.01	.35	.58	.00	.71
3 Boots49	.00	.00	.00	.35	.58	.00	.70
Factor 7: 01:F4 Reasoning (RG); order 1								
8 Figure Changes60	.17	-.02	-.06	.27	.03	.27	.53
7 Mixed Series65	.21	-.02	.07	.26	.00	.25	.60
9 Teams53	.21	.05	.00	.18	-.04	.21	.41
Sum of Squares:	2.84	.61	.54	.80	.48	.67	.18	6.13

d), but the number of factors at the different orders can vary, as long as the number of factors at any given order is less than the number of factors at the next lower order.

The equations shown here assume that variables can be expressed as linear weighted sums of factor scores. This is, of course, only a convenient simplifying assumption, but it appears to be an effective one for data analysis. It implies, however, that the weights assigned to each variable in the factor matrix are the same for all individuals in the sample being studied. That is, it implies that the manner in which abilities determine performance is the same for all individuals, no matter what their individual strategies in approaching test tasks may be. No method of factor analysis (whether "exploratory" or "confirmatory") can reveal differential patterns of weights for different individuals. Investigations of the possibility that individuals differ in their strategies would have to employ special techniques, such as those employed by French (1965), who performed separate factor analyses for groups reporting different strategies, or by MacLeod, Hunt, and Mathews (1978), who were able to identify, by examining fits of data to theoretical models, groups employing different strategies of performance. Even if it is assumed that the parameters of factorial equations vary over individuals, the weights found in factor analysis studies can be taken to be averages of the weights for different individuals.

The equations shown here also assume that test performance can be expressed as a function of series of *orthogonal*, independent factors. To give a simple analogy, consider the typical correlation of about .6 between height and weight of human individuals. The factorial interpretation would show that height is a function of a general "size" factor plus a special independent component for height; similarly, weight is a function of a general size factor plus a special independent component for weight. Furthermore, our equations imply that many psychological ability variables are a function of *several* independent factors of ability. For example, in Table 1 (ignoring loadings near zero), it is seen that the score on the Word Meaning subtest is a function of (a) the individual's standing on a general factor that applies over all tests in the battery, (b) the individual's standing on a factor identified as Gc ("crystallized intelligence") that applies over tests 1 through 4, and (c) the individual's standing on a special "verbal" factor that appears only in the two tests Word Meaning and Odd Words. Some may argue that these factors are merely mathematical artifacts, but I would make the case that the factors are real. Factors are not mere artifacts of the statistical technique: In some sense they are real, just as gravity, mass, distance, and so on, are real variables in the analysis of object motions. The third-order general factor expresses the tendency of the individual to perform at a certain level on *any* cognitive test; the second-order "crystallized intelligence" factor expresses the tendency of the individual's performance on certain kinds of tests to deviate (up or down) from performance expected from the general factor, and the first-order "verbal factor" score expresses the tendency of the individual's performance on verbal tests to deviate (up or down) from what would be expected on the basis of standings on the general factor and on the crystallized intelligence factor.

The use of linear equations implies an assumption about the way in which abilities act in combination to determine test performance. Specifically, it implies that abilities can compensate for each other. That is, for example, if an individual is low on one ability that is important in determining test performance, high ability on another important ability can compensate for this, producing an average score on the test. In the real world, however, abilities may not act in combination in this way. For example, successful performance on a task may require at least a certain minimum level of each of two abilities. The possibility of discovering such a way in which abilities combine is excluded by the linear methodology of factor analysis.

The only additional assumption that is made by the general theory of testing is that the abilities or latent traits measured by psychological tests (or observed by other means) are relatively permanent and stable, in the sense that over long periods of time the individual's standing on a trait changes, if at all, only gradually.

CHARACTERISTICS OF THE DATA

The factor-analytic enterprise, over the years, has become a search for the minimum number of “factors” or sources of variance that are needed to explain (at least statistically) the widest variety of psychological variables that can be established in a particular domain—in the present case, cognitive abilities. The available data thus derive from hypotheses of many individual test constructors as to what kinds of cognitive ability exist and can be measured. I have collected (Carroll, 1993) a large sample of factor-analytic datasets and have attempted to sort out what they show, mainly through reanalysis by a consistent set of procedures. It is from these data that, one could hope, a “theory of intelligence” could be constructed—at least a preliminary description of a state of affairs in terms of a series of factors and the ways in which they may be assumed to influence test scores or other observations of behavior.

Admittedly, this would be only a partial theory of intelligence if it is limited to the analysis of test scores and the like. It would say nothing about how factors arise, or their possible genetic and environmental sources. Other kinds of data would be necessary to round out a theory—data from studies of behavioral genetics (e.g., twin studies), from studies of educational interventions and treatments, from longitudinal studies of ability development, and so on. But at least this partial theory would define and delineate the variables that should be involved and investigated in other studies. For example, it is my impression that studies in behavioral genetics (e.g., Cardon, Fulker, DeFries, & Plomin, 1992) have only begun to consider the many kinds of ability that have been disclosed by factor analysis and to take seriously the information from factor analysis about the structure of abilities.

Limitations of the Data

The data that I have been able to analyze suffer from many limitations.

Size. Most factor studies have been limited to a relatively small number of variables. (In my sample of 461 datasets, the median number of variables studied was about 20, though this number ranged from 5 to 99.) On the assumption that there are actually 40 common factors needed to account for all intellectual performances, and that at least three variables are needed to define each such factor, this means that an ideal factor study would employ at least 120 variables. Assuming that 10 minutes are required to secure a meaningful score on each variable, such a battery would require 1,200 minutes or 20 hours for administration. Even in a military organization, it would be difficult to secure this much testing time. For this reason, the domain of cognitive abilities has had to be studied in a piecemeal fashion.

A related problem concerns the number of subjects required. In my sample of datasets, the number of subjects ranged from about 20 to more than 8,000, with a median of 162. This median is a barely adequate number for statistically stable results, and of course the number of subjects was less than 162—sometimes much less—for half of my datasets.

Poor definition of variables. In any given factor-analytic study, there is little guarantee that the variables employed constitute truly adequate measuring instruments from the standpoint of defining and interpreting the factors that result from analysis. The variables used in many factor-analytic studies are mere collections of items that investigators put together in an effort to measure a particular construct. There is often no assurance that the items measure a unidimensional ability; there is little or no information on how well the items individually perform in measuring whatever is measured, for the items have not usually been subjected to an item analysis. One of the greatest obstacles I have encountered comes from the fact that tests are almost always administered within a time limit such that not every subject is allowed to attempt all items. Under these conditions, it is clear (and known from a number of studies of the matter) that the scores are functions of two types of latent variables—the rate of performance and the accuracy of performance—which in many instances are largely independent. In most factor-analytic studies, it is therefore difficult or impossible to tell the extent to which a factor depends on (a) rate of performance and (b) accuracy or level of mastery. These and other difficulties are compounded by the fact that investigators frequently fail to report critical details about the nature of tasks that are presented and the conditions under which they are given (e.g., the nature of the instructions to subjects, timing, scoring, etc.).

A basic limitation of the studies I have reviewed is that with few exceptions, they are based on data from psychological tests. It has often been argued that psychological tests are atypical measuring instruments, insofar as the psychological test is in many respects an artificial situation. We have few studies in which observations of performance and problem solving in “real-life” situations have been employed. For the most part, we can thus far construct only a theory of intelligence as it is manifested in performances on psychological tests. It is a matter of speculation whether such a theory can be generalized beyond the testing situation. Many kinds of evidence, however, lead me to believe that it can.

LIMITATIONS IN THE ANALYSIS

Analyzing factor-analytic data is in many respects an art rather than a science. There are many problems to be solved, and many options. I believe, neverthe-

less, that the analyses I have made are as sound as they can be made according to currently accepted canons of exploratory factor analysis. This is not the place to describe the techniques of analysis I have used, but a few hints may be given. I have been relatively conservative in deciding on the number of factors to be extracted. Generally, I require iterated principal factor solutions to converge to a very strict criterion (successive communalities must differ by no more than .0005) and after Varimax rotation of the principal factor matrix, each factor must contain salient loadings on at least two variables, where "salient" means that the loading is the highest (in absolute magnitude) for a given variable. The number of factors is selected as the highest number that permits convergence to the stated criterion with adequate salient loadings on factors. Varimax matrices are always considered for oblique rotation, either by the Promax criterion or by procedures recently developed by Tucker and Finkbeiner (1981, 1991). Factorization is carried up to the highest order—usually the second, but occasionally the third—that is viable. The data are subjected to the hierarchical orthogonalization procedure developed by Schmid and Leiman (1957), and factor interpretations are based on the resulting hierarchical factor matrix along with as much information as can be found concerning the nature of the variables. In many instances, factor interpretations are relatively clear, but this is not always the case. Difficulties in factor interpretation are generally traceable to deficiencies in the design of studies or lack of necessary information about the variables.

WHAT THE DATA SHOW

The data reveal what may be called an embarrassment of riches. At one point after I had completed most of the analyses, I took stock of the number of first-, second-, and third-order factors I had derived from 461 data sets:

First-order factors:	2,272
Second-order factors:	542
Third-order factors:	36
<i>Total</i>	2,850

At that point, factors were given tentative interpretations and sorted roughly by domain:

General abilities	459	Visual perception abilities	405
Reasoning abilities	241	Auditory perception abilities	42
Language abilities	367	Numerical facility	81
Memory abilities	251	Mental speed abilities	95

Idea production abilities	252	Interest and motivation factors	12
Attentional abilities	24	Personality and affective factors	17
Interpersonal behavior factors	7	Administrative behavior factors	9
Knowledge factors	72	Education & social background factors	9
School achievement factors	32	Factors whose interpretation was postponed	354
Psychomotor factors	87	Total	2,850
Sensory abilities	34		

Note that these are counts of what I call “token factors” rather than factor types. That is, they are counts of factors identified in 461 datasets. The factors are not all different, of course; the next task was to try to determine how many different factor *types* could be identified. Nevertheless, this count of token factors as sorted into various domains gives an impression of what kinds of variables have been most frequently employed in studies of cognitive abilities. The major categories at the head of the list (general abilities, reasoning abilities, language abilities, etc.) reflect the interest of investigators in studying basic aptitudes. Toward the end of the list, the categories represent domains that are probably not perceived as central to intellectual behavior and performance, but that were included in investigations for various reasons.

The interpretation of a factor can be thought of as a kind of theory construction. As applied to a factor identified in a particular study, it is an effort to develop a statement or series of statements to explain why variables have high salient loadings on the factor, in contrast to variables that have low or vanishing loadings. If factors found in different studies are proposed as being identical, the statements developed to interpret them must be shown to be properly descriptive of all such factors and applicable to explaining the high and low loadings of variables on them. The interpretive statements must also serve to explain why factors that are shown to be factorially distinct from a given factor are theoretically expected to be distinct. It is also desirable that interpretive statements about factors be productive of hypotheses to be tested in further factor-analytic studies.

These comments about factor interpretation apply equally to factors at different orders. At the second order, for example, the goal in interpreting a factor is to develop a statement that explains the variation in factor loadings of first-order factors on that factor. Similarly, a third-order factor is to be interpreted as a theoretical statement explaining variation in factor loadings of second-order factors on that factor. Such statements are to be regarded as applicable to the interpretation of similar factors found in different studies; in this way, a generalization of theory is to be achieved. Unfortunately, at the higher orders of analysis, it is common to find that the number of factors at a lower order is very small, thus limiting the information available for factor interpre-

**Table 2. Loadings of Variables on
Factors CS and CF Illustrative Data from Three Data Sets**

<i>Variable</i>	<i>Factor Loadings on:</i>	
	<i>CS</i>	<i>CF</i>
Dataset BACH21 (Bachor, 1977)		
6 Hidden Figures.....	.50	.00
9 Group Embedded Figures Test.....	.41	.25

3 Figural Intersections Task.....	-.02	.57
11 Stencil Design (Feuerstein Task).....	.00	.50
8 Hidden Patterns (<Thurstone Designs).....	.21	.42
7 Raven Progressive Matrices.....	.03	.41
Dataset BECH01 (Bechtoldt, 1947)		
17 Mutilated Words (Group Test).....	.47	.22
24 5-letter Words (find words with "s").....	.44	-.01
28 Find digit groups with 2, 6, and 9.....	.27	-.09
33 Four-letter Words (find in pied letters).....	.35	.17

43 Shape constancy.....	.00	.31
31 Letter Groups.....	-.01	.30
13 Hidden pictures (group test).....	.25	.28
7 Size Comparison.....	.00	-.41
16 Word Checking.....	.00	-.54
Dataset BOTZ01 (Botzum, 1951)		
31 Four-letter words (find in pied letters).....	.48	.00
27 Street Gestalt.....	.47	-.01
35 Hidden Pictures.....	.43	.02
29 Mutilated Words.....	.38	-.05
30 Incomplete Words.....	.34	.04
28 Backward Writing.....	.33	-.13

21 Gottschaldt Figures.....	.28	.39
22 Designs.....	.36	.34
25 Mechanical Movements.....	.01	.32
26 Hidden Words.....	.18	.20

tation. However, in this case it is often helpful to refer to the loadings of individual variables on a higher-order factor, noting variation in these loadings in terms of the interpretation proposed for the higher-order factor.

These points can be illustrated by recounting experiences in interpreting particular factors. For this purpose I select two first-order factors found in the domain of visual perception, traditionally named Speed of Closure (CS) and Flexibility of Closure (CF) and regarded as distinct. Factors assigned to factor CS were found in 38 datasets in my survey, and factors assigned to CF were found in 21 datasets. In developing interpretations of these factors, all relevant data were considered, but in order to consider the differences between these

factors, the most valuable data were from the 12 datasets in which both CS and CF factors were found. Because of space limitations I can give and discuss here only a small sample of the data available. Table 2 displays that very small sample, from three arbitrarily selected datasets.

For each dataset, the table lists the variables with salient or at least apparently significant loadings on each factor, in the order of the algebraic size of the loadings—first for factor CS and then for factor CF. The names of the variables are given, but these names are often only faintly suggestive of the nature of the underlying tests. Interpretation of factors requires reference to statements in the original studies, or in other sources, for information on exactly what types of tasks were presented to subjects, what types of instructions and fore-exercises were given, how the tests were administered (with or without time limits, individually or in groups, etc.), and how the tests were scored. Such information can be voluminous, but at the same time it can have many gaps because investigators frequently fail to mention important details. For example, in describing the test Hidden Figures, Bachor (1977) stated that a simple design is presented on one side of the page, and a complex design on the other side of the page. He also stated that subjects were asked to find the simple design embedded in the complex design, but he failed to state how the subject was asked to indicate this—was it by overmarking the design, for example?

Because it is impossible to give here complete information on the variables, I can only report that careful consideration of all data available led to the following interpretations of the factors:

Factor CS (*Speed of Closure*). The characteristic process in variables measuring this factor is one of apprehension and free-response naming of a visuospatial form in a visual presentation when the form is presented incompletely or with disguise, degradation, or obscuration. To the extent that a test involves presentation of degraded visual materials, it is likely to have significant loadings on factor CS. Generally the subject is not informed what the form is (i.e., what its name is, or in what category it might be found). Even if the subject is informed about what the form is, or how it might be categorized, this does not necessarily lead to speedy apprehension. Thus there are both speed and level-of-difficulty aspects in the factor. Scores are usually based on speed of response. Various types of forms to be apprehended can be used in tests of this factor. Most often, as in the Street Gestalt Completion or Hidden Pictures test, it is a picture of an object that has been obscured by degradation (omitting parts of the normal stimulus). Or it can be a printed word, as in the Mutilated Words test, where the letters of the word have been partially effaced in a seemingly random way. If the sample contains individuals who have difficulties in reading even normally presented words, a separate Verbal Closure factor may appear for tests involving printed materials, contrasting with the usual Speed of Closure factor that applies mainly to pictorial materials.

Factor CF (Flexibility of Closure). The characteristic process is apprehension and identification of a visuospatial form (generally, a geometric design) when the form is presented (“embedded”) in a background context that obscures the separate figural properties of the form. Generally the subject is shown or told what form is to be looked for. In some tests, there may be a component of visual memory in that the subject must keep in mind precisely what visual form must be looked for, but in other tests the subject is able continually to refer to a presentation of the form to be looked for. The target visual stimulus (that in which a simple form is contained) is usually constructed in such a way that the simple form is obscured by what may be called *geometrical camouflaging*, that is, adding lines in the region contained in and surrounding the simple form. The score on the test is a function of speed of apprehension, but frequently low scores result because subjects are not at all able to identify the simple form in its context within the allotted time. (It may be noted that there is no true “flexibility of closure” element in this factor; a better name and interpretation for the factor would be “speed of detecting and disembedding a known stimulus array from a more complex array.”)

Factors CS and CF are only two of a rather large number of first-order factors that can be identified in the factorial literature. In most cases, theoretical statements as to the nature of these factors, similar to those offered here for factors CS and CF, can be made, although it is evident that much further research is needed to confirm and refine these theoretical statements. It may be helpful to report the theoretical characterizations of the more prominent of these first-order factors, in that they define special aspects of cognitive ability. This is done by domains.

FACTORS IN THE DOMAIN OF LANGUAGE COMPETENCE

Factor LD (Language Development): The extent to which the individual has acquired competence and comprehension in the *spoken* native language with respect to lexicon, syntax, and other aspects of language structure.

Factor V (Printed Language Comprehension): The extent to which the individual has acquired competence and comprehension in the *written* form of the native language.

Factor VL (Lexical Knowledge): The extent to which the individual has acquired (either in spoken or printed form) the lexical wealth of the native language, including meanings of lexical forms.

In addition, various special factors in the language domain can be distinguished, such as RC (Reading Comprehension), RS (Reading Speed), SG (Spelling Ability), PC (Phonetic Coding), MY (Grammatical Sensitivity), LS (Listening Ability), OP (Oral Production and Speech Fluency), and WA (Writing Ability). These pertain to specialized skills that are mainly the result of learning and practice.

FACTORS IN THE DOMAIN OF REASONING AND THINKING

Factor RG (Sequential Reasoning): The ability to proceed correctly through one or more deductive steps in a reasoning task.

Factor I (Induction): When the subject is given appropriate stimulus material, the ability to induce and apply rules and generalizations that govern that material.

Factor RQ (Quantitative Reasoning): The ability to solve problems that call for the use of quantitative concepts.

Factor RP (Piagetian Reasoning): The ability to solve problems involving such “Piagetian” concepts as seriation and conservation. As yet, the status of this factor relative to other reasoning factors is unclear.

FACTORS IN THE DOMAIN OF MEMORY AND LEARNING

Factor MS (Memory Span): The ability to repeat, immediately, a series of auditory or visual stimulus materials. There is some evidence that this depends on two components: (a) registration of the stimulus materials, and (b) memory for the order of the stimuli.

Factor MA (Associative Memory): The ability to form arbitrary associations in stimulus materials such that on testing, the individual can recall what stimulus is paired with another, or recognize, in a series of test stimuli, what stimuli were experienced in a study phase.

Factor M6 (Free Recall Memory): The ability, in a test phase, to recall material presented in a study phase, when the amount of material to be remembered exceeds the individual’s memory span.

Factor MM (Meaningful Memory): Ability to recall, in a test phase, material learned in a study phase, when the material in the study phase is meaningfully connected.

Factor L0 (General Learning Ability): An individual differences parameter that appears to apply to any cognitive learning task. In addition there is evidence of factors of learning ability that are specific to particular kinds of learning tasks.

FACTORS IN THE DOMAIN OF VISUAL PERCEPTION

Factor VZ (Visualization): Ability in manipulating visual patterns, as indicated by level of difficulty and complexity in visual stimulus material that can be handled successfully, without regard to the speed of task solution.

Factor SR (Spatial Relations): Speed in mentally manipulating relatively simple visual patterns, by whatever means (mental rotation, transformation, or otherwise).

Factor CS (Speed of Closure): Speed in apprehending and identifying a visual pattern, without knowing in advance what the pattern is, when the pattern is degraded, disguised, or obscured in some way.

Factor CF (Flexibility of Closure): Speed in finding, apprehending, and identifying a visual pattern, knowing in advance what is to be apprehended, when the pattern is disguised or obscured by a surrounding context. (Factors CS and CF were discussed in more detail earlier.)

Factor P (Perceptual Speed): Speed in finding a known visual pattern, or in accurately comparing one or more patterns, in a visual field such that the patterns are not disguised or obscured.

FACTORS IN THE DOMAIN OF AUDITORY RECEPTION

Factor US (Speech Sound Discrimination): Probably independent of general language competence (factor LD), a special ability to discriminate the phonemes of the native language.

Factors U3 (General Sound Discrimination): Ability to discriminate tones and sequences of tones with respect to basic attributes such as pitch, intensity, duration, and rhythm.

Factor U1 (Auditory Cognitive Relations): Ability to make judgments of complex relations among tonal patterns.

Factor U9 (Musicality Judgments): Ability to make discriminations and judgments of musical materials with respect to expressive aspects, particularly phrasing, tempo, and intensity variation.

Factor UR (Speech Perception Under Distortion): Ability to understand speech that is masked or otherwise distorted.

Factor UK (Temporal Tracking): Ability to track temporal events (usually auditory, but possibly also in other modalities) on a short-term basis so as to count or rearrange them.

Factor UM (Auditory Memory): Ability to retain, at least on a short-term basis, images of auditory events such as tones, tonal patterns, and voices.

Factor UP (Absolute Pitch Ability): Long-term memory for the categories of musical pitch. (This factor has not been identified in factor-analytic studies but there is abundant evidence for it as a dimension of individual differences, with various special properties—rarity and its all-or-none character.)

FACTORS IN THE DOMAIN OF IDEA PRODUCTION

Factor FI (Ideational Fluency): Speed and success in thinking of and reporting (usually in writing) a series of different verbal responses falling in a specified semantic class.

Factor NA (Naming Facility): Speed and success in evoking and reporting (orally or in writing) an accepted name for a given thing, as cued by the thing itself or a picture of it, or in some other appropriate way.

Factor FA (Associational Fluency): Speed and success in thinking of and reporting (usually in writing) a series of different verbal responses that are semantically associated with a given stimulus.

Factor FE (Expressional Fluency): Speed and success in thinking of and reporting (usually in writing) a series of syntactically coherent verbal responses under highly general or more specific cueing conditions.

Factor FW (Word Fluency): Speed and success in thinking of and reporting (usually in writing) one or more language units (usually, words) that have specified phonemic or graphemic properties. The factor is also measured by tasks (e.g., anagrams) that indirectly involve this kind of language unit evocation in their solution.

Factor SP (Sensitivity to Problems): Speed and success in evoking and reporting (usually in writing) solutions to “practical, real-world” problems, or novel ways of using objects and materials.

Factor FO (Originality/Creativity): Speed and success in evoking and reporting (usually in writing) unusual or original verbal/ideational responses to specified tasks.

Factor FF (Figural Fluency): Speed and success in producing (usually by drawing) a variety of “figural” responses to specified tasks.

Factor FX (Figural Flexibility): Speed and success in dealing with figural tasks that require a variety of approaches to a solution.

FACTORS IN THE DOMAIN OF COGNITIVE SPEED

Although many of the factors already listed involve speed, the following may also be recognized:

Factor N (Numerical Facility): The ability to perform simple arithmetical operations quickly and accurately.

Factor R9 (Rate of Test Taking): Speed in working through and performing simple cognitive tests.

In addition, a number of factors can be measured with reaction times in various kinds of *elementary cognitive tests* (ECTs). As yet, evidence is unclear as to exactly what these factors are, or as to how they are differentiated. A tentative listing of these factors is as follows:

Factor R1 (Simple Reaction Time): Speed in responding to the onset of a stimulus, or at least leaving a “home position” in order to make a response.

Factor R2 (Choice Reaction Time): Speed in responding to one of two or more possible stimuli, or at least leaving a “home position” in order to indicate the response.

Factor R3 (Movement Time): Speed in moving from a home position to a response key, in paradigms in which such movements are required and operationally distinguished from decision times. This is probably *not* a measure of any kind of cognitive speed; it is better considered as a psychomotor factor (speed of limb movement).

Factor R4 (Speed of Semantic Processing): Reaction times for ECTs in which the decisions to be made by the subject require encoding and mental manipulation of stimulus content.

Factor R5 (Slope Parameters, Visual and/or Memory Search): Rate of decision time as a function of the number of elements involved in the search task.

Factor R6 (RT or Intercept Parameters, Visual and/or Memory Search): Reaction time in ECTs involving visual and/or memory search.

Factor R7 (Speed of Mental Comparison): Reaction time measured in tasks in which stimuli must be compared for particular attributes.

FACTORS OF KNOWLEDGE

Various factors of knowledge can be identified in the factorial literature, for example, knowledge of mathematics, general science information, knowledge of mechanical principles, knowledge of the functions of tools and various machines, and so on, but it is evident that they represent simply the degree to which an individual's knowledge is specialized in a particular field. Although such factors tend to show correlations with certain factors of cognitive ability, the structuring of knowledge factors is not viewed as contributing to the structuring of intelligence or cognitive abilities, and for this reason they are not presented or discussed here.

One interesting and possibly important point may be mentioned, however. This is that tests of general vocabulary (measuring factor VL) tend to be highly correlated with tests of "general information" of the type measured, say, by the Information subtests of the Wechsler Adult Intelligence Scale (WAIS). This suggests that general information of this kind is acquired in the same way as advanced vocabulary knowledge.

HIGHER-ORDER FACTORS

After factor rotation to simple structure, first-order factors are often found to be correlated. The correlation matrix of a set of first-order factors can be factored to reveal one or more second-order factors. If after simple structure rotation these second-order factors are found to be correlated, their matrix of correlations can in turn be factored to reveal one or more third-order factors. In my datasets, more than one third-order factor was never found, even in one of

the largest datasets—that by Hakstian and Cattell (1974) in which, for a correlation matrix of 57 variables, 19 first-order factors, five second-order factors, and one third-order factor were found. As already noted, a large number of second-order factors were found in my datasets, and 36 instances of a third-order factor.

In various writings, Hakstian, Horn, and Cattell (e.g., Hakstian & Cattell, 1978; Horn, 1988) have proposed a set of second-order (or “second-stratum”) abilities that account for the interrelations of first-order abilities. In general, my reanalyses tend to confirm the existence of these abilities, named, symbolized, and interpreted as follows:

Factor Gf (Fluid Intelligence): A factor embracing basic capacities to perform intellectual tasks that only minimally involve the use of acculturated knowledge, such as reasoning, induction, and spatial visualization. (The term “fluid,” proposed by Cattell, 1943, was intended to suggest that factor Gf “flows into” a wide variety of intellectual performances.)

Factor Gc (Crystallized Intelligence): A factor that represents the degree to which the individual has been able to utilize or “invest” basic intelligence in acquiring various types of acculturated knowledge, such as knowledge of and competence in the native language, and knowledge of quantitative concepts and the ability to solve problems involving quantitative concepts. It thus embraces such first-order factors as V (Verbal Ability), RQ (Quantitative Reasoning), and MK (Mechanical Knowledge).

Factor Gv (Visualization Capacity): A factor that represents a general capacity to apprehend and visualize spatial forms, and thus presumably embraces the various first-order factors found in the domain of visual perception, such as VZ (Visualization), SR (Spatial Relations), CS (Speed of Closure), CF (Flexibility of Closure), and P (Perceptual Speed).

Factor Gs (General Cognitive Speed): A factor presumably encompassing or influencing any ability that involves speed of cognitive activity or performance.

Factor Gm (General Memory): The general capacity to commit material to memory, entering into such factors as MS (Memory Span), MA (Associative Memory), and MM (Meaningful Memory).

Factor Gr (General Retrieval Capacity): General ability to retrieve material rapidly from memory storage, as shown in factors involving production of ideas.

Factor Ga (General Auditory Perception Capacity): A factor that influences, in common, the various first-order factors in the auditory perception domain.

Numerous instances of most of these factors were found in the datasets I surveyed and reanalyzed. Only factor Ga was poorly represented, mainly because relatively few datasets have focused on auditory abilities or disclosed more than one auditory ability factor at the first order.

Of considerable interest was whether the postulated Gf and Gc factors would appear as distinct second-order factors. The data are not entirely clear

on this matter, because although second-order factors appeared that seemed interpretable as Gf and Gc, respectively, the first-order factors subsumed by them were somewhat variable. Also, there is a self-fulfilling prophecy problem in separating and classifying these factors. That is, given two factors, one of which was loaded with first-order factors in the language or idea production domain, and the other of which was loaded with first-order factors in the reasoning and/or visual perception domain, the temptation was to interpret the first as factor Gc and the second as Gf, automatically producing a differentiation between the factors when other assignments or interpretations might have been more appropriate. Nevertheless, my impression is that my data tend to support the existence of second-stratum factors Gf and Gc in the forms proposed by Horn, Cattell, and others.

The data also were found to support the existence of a third-order general factor that enters into most factors of cognitive ability. This was clearly found, of course, only in the relatively few datasets that exhibited two or more second-order factors. The best data came from reanalysis of dataset HAKS01 (Hakstian & Cattell, 1974) as mentioned previously, showing five second-order factors. The interpretations of these factors, and their loadings on the third-order factor, were as follows:

Factor Gm (General Memory)	.76
Factor Gf (Fluid Intelligence, including Visualization)	.62
Factor Gc (Crystallized Intelligence)	.52
Factor Gr (General Retrieval)	.22
A general information factor	.21

It is to be noted that factors Gm, Gf, and Gc showed the highest loadings on the third-order general factor. Performance on tests with high loadings on factor Gr would be little affected by standing on the general intelligence factor.

Similar data from one other dataset are of interest. This was a dataset studied by Gustafsson (1984) in support of his HILI model of intelligence, in which he found, using LISREL structural modeling techniques, that the third-order general intelligence factor was identical to the second-order Gf factor. According to my exploratory factor analysis of his data, however, the third-order general factor is closer to a second-order Gv (General Visualization) factor than to Gf. The loadings of my three second-order factors on the third-order factor were as follows:

Factor Gv (General Visualization)	.93
Factor Gc (Crystallized Intelligence)	.63
Factor Gf (Fluid Intelligence)	.56

Although it is currently fashionable to favor structural modeling analyses, there is something to be said for exploratory techniques such as what I have used, in that they “let the data speak for themselves” to a greater extent than do structural modeling techniques, which are amenable to rather wide variations in types of hypotheses that can be tested, and thus may be subject to investigators’ biases and preferences for particular hypotheses.

Many of the datasets I surveyed failed to disclose distinct Gf and Gc factors; rather, they tended to yield a single second-order factor that could be interpreted as, perhaps, a combination of Gf and Gc, or as one of the other proposed second-order factors. This result could often be attributed to study designs that were not adequate for studying the structure of higher-order abilities, in that the variety of variables was not sufficient to yield more than one second-order factor.

THE THEORY SO FAR

We have here, I believe, the makings of a theory of cognitive abilities. (I hesitate to say that we have a theory of “intelligence” because it is evident that intelligence is not one thing, but many.) I call this a *three-stratum* theory because it postulates factors at three hierarchically arranged strata.

First of all, we have evidence that a fairly substantial number of separate cognitive abilities exist. I have mentioned some 40 or more distinguishable abilities represented by first-order factors that can be identified. For each such ability, it has been possible to develop statements about the characteristics of tests and tasks measuring that ability, and these statements constitute, as it were, a *theory* of that ability. Some abilities are more satisfactorily defined than others, but for all abilities, further research is possible to refine the theoretical statements thus far achieved.

Second, we have evidence that broader, higher-order abilities exist, and theoretical statements can be made about the nature and generality of these abilities, and the kinds of first-order abilities they are likely to embrace. These theoretical statements should also make it possible to predict, from characteristics of tasks, what higher-order factors are likely to come into play in the performance of those tasks.

One higher-order factor appears to be of particular importance—namely, the *g* factor that has been postulated ever since Spearman’s time. In the case of datasets that have been adequately designed to reveal it, it shows up as a third-order factor. Some insight into its nature can be gained by considering what loadings different second-order factors have on it. Earlier I reported results from two studies. In one case, general memory, fluid intelligence, and crystallized intelligence had the higher loadings, in contrast to lower loadings for a

general retrieval factor and a general information factor. In the other case, the highest loading was for general visualization, with somewhat lower loadings for crystallized and fluid intelligence. Unfortunately, these two studies are among the very few that were adequately designed to yield reliable information on the factorial composition of the general intelligence factor. Pending further research, one can only say that the general factor appears to have its highest loadings for factors and variables that involve the *level of complexity* at which individuals are able to handle basic processes of induction, deduction, and comprehension. I do not believe that the general factor pertains in any essential way to the speed with which individuals handle these processes. As I read the evidence thus far, cognitive speed has only a very low correlation with general intelligence, if any at all.

In my view, the results of the factor-analytic investigations I have surveyed, together with the theoretical statements that can be made and confirmed about the nature of the ability factors that have been identified, constitute a highly articulated theory of cognitive abilities. Essential to the theory would be good estimates of the parameters of the theory—that is, the weights in its factorial equations. Also, the theory should be buttressed by information as to the developmental and life-span aspects of the individual differences parameters in the theory.

It would be going far beyond this theory to speculate about the extent to which factors are affected by genetic and environmental sources of variance, and I shall not attempt to do so at this time. Nevertheless, I see the theory developed here as providing guidelines for the variables to be investigated in genetic and environmental studies.

REFERENCES

- Bachor, D. G. (1977). *Information processing capacity and teachability of low achieving students*. Unpublished doctoral dissertation, University of Toronto, Canada (Canadian Theses on Microfiche 30331).
- Bechtoldt, H. P. (1947). *Factorial study of perceptual speed*. Unpublished doctoral thesis, University of Chicago, IL.
- Botzum, W. A. (1951). A factorial study of the reasoning and closure factors. *Psychometrika*, 16, 361–386.
- 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.
- Carroll, J. B. (1987). New perspectives in the analysis of abilities. In R. R. Ronning, J. A. Glover, J. C. Conoley, & J. C. Witt (Eds.), *The influence of cognitive psychology on testing* (pp. 267–284). Hillsdale, NJ: Erlbaum.

- Carroll, J. B. (1990). Estimating item and ability parameters in homogeneous tests with the person characteristic function. *Applied Psychological Measurement, 14*, 109–125.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Carroll, J. B., Meade, A., & Johnson, E. S. (1991). Test analysis with the person characteristic function: Implications for defining abilities. In R. E. Snow & D. E. Wiley (Eds.), *Improving inquiry in social science: A volume in honor of Lee J. Cronbach* (pp. 109–143). Hillsdale, NJ: Erlbaum.
- Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin, 40*, 153–193.
- French, J. W. (1965). The relationship of problem-solving styles to the factor composition of tests. *Educational and Psychological Measurement, 25*, 9–28.
- Gustafsson, J.-E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence, 8*, 179–203.
- Hakstian, A. R., & Cattell, R. B. (1974). The checking of primary ability structure on a broader basis of performances. *British Journal of Educational Psychology, 44*, 140–154.
- Hakstian, A. R., & Cattell, R. B. (1978). Higher-stratum ability structures on a basis of twenty primary abilities. *Journal of Educational Psychology, 70*, 657–669.
- Horn, J. L. (1988). Thinking about human abilities. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., pp. 645–685). New York: Plenum.
- MacLeod, C. W., Hunt, E. B., & Mathews, N. N. (1978). Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning and Verbal Behavior, 17*, 493–507.
- Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22*, 53–61.
- Schutz, R. E. (1958). Factorial validity of the Holzinger-Crowder Uni-Factor Tests. *Educational and Psychological Measurement, 18*, 873–875.
- Tucker, L. R., & Finkbeiner, C. T. (1981). *Transformation of factors by artificial personal probability functions* (Res. Rep. RR-81-58). Princeton, NJ: Educational Testing Service.
- Tucker, L. R., & Finkbeiner, C. T. (1991). *Transformation of factors by artificial personal probability functions*. Unpublished manuscript.

Chapter 4

***The Bio-Ecological
Theory of Intelligence:
A Developmental-Contextual
Perspective***

Stephen J. Ceci Maggie Bruck

Cornell University McGill University

Since the early 1980s, a great deal of new psychometric evidence has been amassed to support a hierarchical model of intelligence, with the most general intellectual factor, denoted g , located at the apex of a pyramid that subsumes family and specific factors. The construct of g not only includes the assumption of a general form of intelligence but also the assumption that this intelligence is biologically based and genetically determined. Based on a review of a vast literature, one might conclude that there is overwhelming evidence to support this view of a general, biologically based intelligence. It is not, however, accepted by all researchers. The purpose of this chapter is to outline an alternative account.

At the outset, let us be clear about our biases: We are not psychometric researchers, but developmentalists. In general, developmentalists regard psychometric claims about *g* as irrelevant or superfluous to developmental accounts of cognitive growth despite the fact that few of us have actually ventured into the psychometric field. This is a pity because developmentalists' ignorance about psychometrics can lead us to dismiss carefully reasoned and well-documented claims about *g*. Furthermore, regardless of the ultimate validity of these claims, they require the attention and understanding of all social scientists as they have extreme practical importance; they reflect the current basis by which America chooses its elite for education, training, and work. That is, many of the screening measures used in education, government, and industry are selected because they are purportedly good measures of *g*, presumed to be a general, biologically based index of intelligence. Having said this, however, there are two important respects in which psychometric research is not as pertinent to developmentalists as it might be. First, dominant psychometric models of intelligence do not make explicit the psychological mechanisms that translate genes into cognitive phenotypes. Second, these models do not focus on transition mechanisms, or provide explanations of how the organism's cognitive system unfolds over time, other than providing descriptive reports of changing h^2 or *g* with age (e.g., Plomin, 1985). In short, to developmentalists, the field of psychometrics tends to be less "psycho" than "metric."

In this chapter, we shall first describe the major assumptions of the traditional psychometric model. We shall then provide several direct attacks on key psychometric concepts, such as the heritability and the presumed generality of intelligence, or *g*. Following this, we will outline the bio-ecological theory of intellectual development, and describe where it is positioned within both the biological and developmental sciences.

ASSUMPTIONS OF TRADITIONAL PSYCHOMETRIC MODELS

The traditional psychometric models (e.g., Humphreys, 1962, 1979; Jensen, 1979; Marshalek, Lohman, & Snow, 1983; Spearman, 1904) contain three assumptions. First, there is a singular, pervasive ability called general intelligence, *g*. The evidence for this assumption is that whenever a battery of diverse mental tests are administered to a diverse sample of subjects, and their scores are intercorrelated, the result is a positive manifold of correlations. That is, the performance of the individual tends to be consistent across tests. The assumption is that consistent performance on such seemingly different tests such as verbal reasoning, cultural knowledge, and spatial abilities can only reflect the fact that each test is saturated with a common factor, namely general intelligence or *g*.

The second assumption of traditional psychometric models is that this general intelligence is biologically based. Many researchers have reached this conclusion because measures of g are correlated with h^2 and other constructs of heritability, as well as with a variety of physiological measures such as cranial blood flow, evoked potential recordings, central nerve conductance velocity, and oscillation (e.g., Eysenck, 1988; Hendrickson & Hendrickson, 1982; Jensen, 1992, in press; Schafer, 1987).

The third assumption of traditional psychometric models is that IQ tests are good measures of this biologically based, general intelligence that permeates most intellectual endeavors, from verbal reasoning, reading comprehension, and cultural knowledge to quantitative, spatial, and mechanical abilities. Construct validation research has attempted to show that IQ tests are more potent predictors of a range of outcomes such as job success, school grades, and training scores than are measures of specific cognitive abilities, motivation, or relevant background experience (e.g., Barrett & Depinet, 1991; Gottfredson, 1986; Hunter & Schmidt, 1982; Hunter, Schmidt, & Rauschenberger, 1984). It is argued that there are lower correlations with these specific factors because these are less perfect markers of g than are IQ tests, which are alleged to be highly saturated with g .

Taken together, the above three categories of evidence have persuaded many to view psychometric test scores as reflections of a singular, biologically based resource pool that permeates virtually all intellectual feats and that is responsible for individual differences in many real-world outcomes. As readers of the journal *Intelligence* know, there is an enormous research literature that is in accord with the predictions of this view, such as the relationship between task complexity and g , the relationship between h^2 and g , and so on. Those who support this view do not deny that the environment plays some role in performance on g -based tests, but they do underscore the relative ascendancy of biology, in some cases pointing out that environments themselves are genetically loaded. That is, the environment is an important contributor to individual and developmental differences by exerting its influence through genetically determined forces, through gene-environment correlations and/or interactions (Plomin & Bergeman, 1991; Plomin & Neiderhiser, 1992):

Labelling a measure "environmental" does not make it a pure measure of the environment. People make their own environments.... The ways in which people interact with their environments—their experiences—are influenced by genetic differences. (Plomin & Neiderhiser, 1992, p. 163)

Scarr's work (1992) provides a paradigmatic example of the biologically oriented interactionists' claim that environmental effects are mediated by genetic influences. Scarr reanalyzed data that had been commonly cited as evi-

dence for large environmental influences on children's IQ. On the basis of these reanalyses, she argued that the distal cause of variation in children's IQs was not the putative differences in a child's proximal environment (e.g., in parental disciplinary styles) but rather variations in mothers' IQs, a presumed genetic marker. Scarr (1989, 1992) showed that the relationship between a maternal disciplinary style and her child's IQ could be accounted for by the fact that mothers who employed positive types of discipline had higher WAIS vocabulary scores than did other mothers. Thus, what appeared to be an environmental influence (type of discipline) on their offspring's IQ was really the result of a gene-environment correlation; the driving force behind such ostensibly environmental variation (related to parental disciplinary style) is, in actuality, genetics. Scarr (1992) asserted that as long as some minimum (nonabusive) environmental context is provided, variations in children's intelligence will be unrelated to variations in their environment:

Being reared within one family, rather than another, within the range of families sampled makes few differences in children's personality and intellectual development. These data suggest that environments most parents provide for their children have few differential effects on the offspring [p. 5].... The important point here is that variations among environments that support normal human development are not very important as determinants of variations in children's outcomes [p. 6].... For children whose development is on a predictable but undesirable trajectory and whose parents are providing a supportive environment, interventions have only temporary and limited effects. Should we be surprised? Feeding a well-nourished but short child more and more will not give him the stature of a basketball player. Feeding a below-average intellect more and more information will not make her brilliant.... The child with a below-average intellect... may gain some specific skills and helpful knowledge of how to behave in specific situations, but (her) enduring intellectual and personality characteristics will not be fundamentally changed. (Scarr, 1992, pp. 16-17)

Taking this argument to the next stage, supporters of the psychometric position attempt to estimate through heritability formulae (see Ceci, 1990a, for a description and interpretation of these estimates) the degree to which genetics and environment influence mental development. Reports of very high heritabilities¹ are now commonplace for cognitive performance. For example, the psychometrist Carroll (1992) has recently claimed that at least 50% of the variance in cognitive ability is genetic in origin, and Plomin and his colleagues have reported heritability estimates for specific groups of individuals that range in excess of .80 (Pedersen, Plomin, Nesselroade, & McClearn, 1992). If these claims are correct, then one might question how much the developing organism's ecology can influence mental development, as at least half, and possibly upwards of 80%, of the variation in intelligence test scores would seem to be accounted for solely by genetic variation.

TAKING ISSUE WITH SOME PSYCHOMETRIC CLAIMS

There is now some evidence to support an alternative interpretation of the psychometric evidence, one that does not require the conclusions above. We will first present a glimpse of this evidence as it bears on two claims, namely the generality of intelligence, that is, g , and its presumed biological basis. We shall argue that the generality of intelligence has been overstated by focusing on analyses that were not designed to probe the possibility that multiple shared cognitive resources are involved in so-called general intelligence, and that g is distributed disproportionately among the lower end of the scale, that is, among persons with low IQs.

The Issue of Generality

Elsewhere, we have suggested several explanations for the correlations among diverse cognitive tests, none of which invoke the concepts of saturation or loading on a general factor, g (Ceci, 1990a, 1990b). First, factor-analytic studies of mental test performance may yield a large measure of general intelligence because they rely on tasks that involve the operation of *many* independent cognitive abilities that interact with each other for successful performance (for a cogent argument see Detterman, 1986). For example, suppose that a battery of three tests (arithmetic, vocabulary, and spatial reasoning) is administered to children and that their scores are factor analyzed and g is extracted. Suppose also, that the arithmetic performance depends on just three cognitive components a (verbal encoding), b (a spatial mapping skill that is relevant for some of the geometric problems), and c (a highly specific quantitative skill that is useful for a broad array of arithmetic problems, but not for nonarithmetical problems). Now, suppose that the two other tests (vocabulary and spatial reasoning) also sample some, but not all, of these same microlevel cognitive components, in addition to some components that are highly specific to themselves. For example, maybe vocabulary requires a (verbal encoding) as well as d , which is the ability to “compare representations” and e , which is a highly specific vocabulary skill. Finally, suppose that spatial reasoning requires b , in order to engage in spatial mapping, d , in order to compare representations, and f , which is a highly specific spatial skill. Then vocabulary and arithmetic performances might be correlated because they share verbal encoding resources (a), arithmetic and spatial reasoning might be correlated because they share a mapping resource (b), and vocabulary and spatial reasoning might be correlated because they share the ability to compare representations (d). This, of course, is a highly contrived example, but it does illustrate the point about the sampling of partially shared components. In principle, g could end up being substantial in magnitude without actually representing any single source of pro-

cessing variance that is common to all three tasks. If we replaced this contrived example with a more realistic one, involving dozens of component cognitive skills, differentially sampled by various tasks, then the possibility that there is a series of partially shared components that give rise to g in the absence of a single pervasively shared cognitive resource becomes not so far-fetched. Recently, Detterman, Mayer, Caruso, Legree, Connors, and Taylor (in press) have reported evidence consistent with this interpretation.

Because the technique of factor analysis can yield a “general” factor that correlates with IQ, some psychometric theorists conclude that this correlation provides validation that IQ tests measure a truly singular yet general factor. As Detterman (1986) points out, however, this type of reasoning is circular at best. One could just as easily have posited that IQ was itself an amalgam of some of the same components that were sampled by the individual tests in the original battery. If this were done, then the IQ test could have been included in that battery to begin with and it would be seen as a marker for a set of separable cognitive factors rather than a marker for a truly general factor.

Rarely do factor analysts attempt to separate such diverse components of cognitive components, structures, and processes involved in the successful completion of psychometric tests. But when this is done (usually by altering the order in which regression analyses are run to see whether the variance associated with one is absorbed by another when it is entered earlier), the evidence in favor of generality is often diminished (Detterman, 1986), and the evidence in favor of a multiplicity of cognitive abilities that are independent of g is compelling, at least to some (e.g., Horn, 1989).

Furthermore, even when generality across tests *is* observed, this occurs disproportionately as a result of persons with very low IQs behaving generally. That is, over most of the IQ range there is a substantially greater differentiation of cognitive abilities (with a resultant lower g) than there is among persons with IQs one or more standard deviations below the mean (Detterman, 1991; Detterman & Daniel, 1989; Detterman & Persanyi, 1990). The first person to recognize that g was smaller among persons with average or above average cognitive ability than among persons with low cognitive ability was Spearman himself (see Deary & Pagliari, 1991). This means that the concept of “general intelligence” that has been at the core of 20th-century thinking about intelligence, is possibly a misnomer; individuals who behave most generally are those with low IQs. If true, then what is “general” is not intelligence but the lack thereof, namely stupidity (Detterman, 1991).

Another reason that g may emerge in so many statistical analyses concerns the fact that some of the cognitive tests included in a battery of psychometric tests require a common degree of elaborateness of the underlying mental representation that is instantiated by these different tests. For example, if encoding, short-term memory, and verbal analogy tasks all rely on the same mental

representation of a semantic field, then they will all be similarly facilitated or limited by the extent to which the mental representation in that field is well elaborated. To take a simple case, if the digit 9 is an unelaborated mental representation which only entails only a few dimensions (e.g., less than 10, but more than 8; odd number), then all cognitive performances that involve this digit will be limited vis-à-vis tasks that entail digits with more elaborate representations. For instance, the speed with which a digit can be recognized when it is followed by a pattern mask will be quicker if its mental representation includes such dimensions as cardinality and root properties, than if it does not. Similarly, the ability to recall the digit will be facilitated by increased elaboration of its representation, and the ability to reason with the digit will also be facilitated by increasing the elaborateness of its representation (i.e., using it to solve an analogy or to fill in a number series).

Thus, intercorrelations among test scores, which are the basis for extracting *g*, can arise from mechanisms that may have little to do with the existence of a singular, underlying resource pool, but may reflect the accumulation and organization of declarative knowledge that leads to mental representations that are shared among some tests as well as the components that are shared across some tasks. Accordingly, one might expect inconsistent performance across tasks that require the same component skill (e.g., encoding, retrieval, verbal rehearsal) whenever the tasks instantiate different knowledge bases. In fact, there are low correlations among encoding tasks when the elaborateness of the underlying mental representation varies across tasks (Ceci, 1990b). Thus, some subjects who are slowest encoding some stimuli are fastest encoding other stimuli. So, it is not easy to argue that those with the least elaborate knowledge representations also have the least amount of *g* because these same individuals are actually superior on tasks which sample domains more related to their knowledge base.

The results of studies that compare specific cognitive skills of experts and novices provide even stronger evidence that the efficiency of information processing is domain specific. These studies indicate that traditional information-processing measures sometimes cannot predict performance in nonacademic domains such as gambling, stock market analysis, or the comprehension of sporting events (e.g., Ceci & Liker, 1986; Ceci & Ruiz, 1992b; Schneider, Körkel, & Weinert, 1989; Walker, 1987). Critics have often dismissed the importance of these results by claiming that the tasks (for example, handicapping races) are highly specific (see Brody, 1992). Although it is certainly true that they are specific, it is precisely the specific type of reasoning that is of interest because it involves highly complex forms of reasoning that can be broken down and studied experimentally. This is an important characteristic that is usually missing from the tests and tasks that have been used in psychometric research. Consequently, the expertise studies possess a high degree of face

validity: To do well on them, one must engage in sophisticated, generative, and novel reasoning—the type that is used to successfully solve some types of important everyday problems.

For example, the study of expert racetrack handicappers revealed that highly complex forms of interactive reasoning, which were independent of IQ scores, formed the basis of a small part of these experts' judgments (Ceci & Liker, 1986). This involved the simultaneous consideration of a multitude of variables, their differential weighting, and the use of an implicit algorithm to combine these weighted variables. This type of reasoning is as complex as any carried out in academia, and it seems to epitomize what is meant by the term *intelligence*. At the very least, it seems closer to what is meant by the term *intelligence* than defining vocabulary words such as *espionage* or answering questions based on the acquisition of cultural knowledge (e.g., “Why do we have child labor laws?”)—items commonly found to be *g*-loaded. Therefore, even though these measures of IQ-independent expertise are specific, they have the advantage of possessing the very qualities we regard as intelligent.

Although they are highly complex, the types of reasoning exhibited by experts is highly content- and domain-specific. For example, expert handicappers do not recognize the identical algorithm and weighting scheme of a stock market simulation that is isomorphic to the racetrack problem (Ceci & Ruiz, 1993). The question such research raises, of course, is why they cannot exhibit similarly complex forms of reasoning in contexts other than the racetrack. Whatever the answer, it appears to be a condition that is pervasive, according to Detterman (1993):

The amazing thing about all of these (reasoning and problem-solving) studies is not that they don't produce transfer. The surprise is the extent of similarity it is possible to have between two problems without subjects realizing that the two situations are identical and require the same process. (Detterman, 1992, p. 13)

While we cannot as yet supply the answer to the above question, we can rule out the factor of IQ. That is, experts of all IQ levels, even those in the superior range, exhibit domain-specific reasoning (Ceci & Ruiz, 1993). It is surprising to see just how context-specific subjects of all IQ levels are, including college-educated ones with presumably above-average IQs. So, the answer to the question of why people can reason more complexly in some domains than others does not appear to be found in their performance on *g*-based tests.

The Issue of Biology

Recently, some of us have suggested that the evidence recruited to support biological interpretations of *g* is open to alternate explanations (Bronfenbrenner, Ceci, & Lenzenweger, 1993). We have argued, for example, that the interpre-

tation given to h^2 is misleading, because h^2 as it is commonly computed (e.g., see Cavalli-Sforza & Bodmer, 1971) can only refer to the proportion of "actualized" genetic variance. Thus, the amount of genetic variance that is unactualized as a result of insufficiencies in the developing organism's ecology is unknown. This means that even if $h^2 = .80$ for a cognitive measure, then it is incorrect to infer that the environment can, at best, influence only 20% of the individual differences on this measure because we have no way of knowing how much unactualized potential exists. In the next section, which outlines the bio-ecological view of intellectual development, we show how traditional measures of h^2 may be highly labile because of the failure to take into consideration a variety of assumptions involving environmental factors which could either increase or decrease heritability estimates.

Before describing the bio-ecological theory of intellectual development, a caveat is in order: None of the foregoing is meant to claim that traditional assumptions about *generality*, *heritability*, and *singularity* are decidedly incorrect. Rather, the claim is that the evidentiary basis which underlies traditional assumptions about them is open to different interpretations; that is, both psychometric test performance and intellectual development result from a concatenation of cognitive, social, and biological factors (e.g., Sternberg, 1985, 1990).

THE BIO-ECOLOGICAL THEORY OF INTELLECTUAL DEVELOPMENT

We turn first to a discussion of some major characteristics of the bio-ecological model. According to the bio-ecological view, the data of intellectual development necessitate a four-pronged framework for their explication, namely, (a) the existence of not one type of intellectual resource, but multiple, statistically independent resource pools, (b) the interactive and synergistic effect of gene-environment developments, (c) the role of specific types of contexts (e.g., interactions, called "proximal processes," as well as more distal environmental resources such as family educational level) that influence how much of a genotype gets actualized in what type of an environment, and (d) the role of motivation in determining how much one's context aids in the actualization of one's potential. We now review these four prongs, and in doing so describe in greater detail how the bio-ecological model accounts for the bedrock data that underpin traditional models.

First, based on the research and arguments presented in the previous section, intelligence is viewed as a multiple resource system. Making this assumption gets around the thorny problem of domain specificity and low cross-task correlations when the same cognitive operation is involved. It also accords with the analyses by Detterman and his colleagues, showing that independent cognitive processes make unique predictions to *g*-based measures.

Second, the bio-ecological view is inherently developmental and interactionist. Like all interactionist perspectives, the bio-ecological view asserts that from the very beginning of life there is an interplay between biological potentials and environmental forces. In order to understand how individuals could begin life possessing comparable intellectual potentials but differ in the level of intelligence they subsequently manifest, the bio-ecological view posits an interaction between various biologically influenced cognitive potentials, such as the capacity to store, scan, and retrieve information, and the ecological contexts that are relevant for each of their unfoldings. At each point in development, the interplay between biology and ecology results in changes that may themselves produce other changes until a full cascading of effects is set in motion.

Although biology and ecology are interwoven into an indivisible whole, their relationship is continually changing, and with each change a new set of possibilities is set in motion until soon even small changes produce large effects. Hence, developmental change is not always or even usually linear, but rather is synergistic and nonadditive. A small environmental influence on a protein-fixing gene may initially result in only tiny changes, but over time the chain of events may produce a magnification of effects on other processes. In addition, certain epochs in development can be thought of as sensitive periods during which a unique disposition exists for a specific cognitive muscle to crystallize in response to its interaction with the environment. During such periods, neurons within specific compartments rapidly "overarborize" (spreading their tentacle-like synaptic connections to other neurons). Even though some of the arboreal connections laid down during these periods of brain spurts will not be used at that time, they can be recruited to enable future behaviors to occur, provided they are not "pruned" because of atrophy or disuse. Siegler (1989) concluded that "the timing of the sensitive period seems to be a function of both when synaptic overproduction occurs and when the organism receives relevant experience" (p. 358). It appears that while some neural processes are more fully under maturational control, others are responsive to the environment, and synapses are formed in response to learning that may vary widely among humans. Similar contextual roles have also been found in the case of various animals' cognitive skills (e.g., Lickliter & Hellewell, 1992; Smith & Spear, 1978).

Third, the bio-ecological model stipulates that "proximal processes" which depend in part on distal resources in the child's environment are the engines of intellectual development. Proximal processes involve reciprocal interactions between the developing child and other persons, objects, and symbols in its immediate setting. In order to qualify as "proximal processes," these interactions must be enduring and lead to progressively more complex forms of behavior. Appropriate proximal processes differ as a function of the developmental status of the organism (e.g., during infancy, a proximal process might be an activity between a caregiver and an infant that serves to maintain the

infant's attention or encourage her to slightly exceed her proximal zone of potential). As we argue later, proximal processes are the engines that drive development; they are the mechanisms that translate genes into phenotypes (Bronfenbrenner, Lenzenweger, & Ceci, in press).

Based on the reanalyses of two different data sets, Bronfenbrenner et al. (in press) have argued that when proximal processes are at high levels in a child's environment, then heritability estimates are high, and yet at the same time individual differences *may* be attenuated. The first part of this claim is no different from that of many in the psychometric community (e.g., Herrnstein, 1973; Humphreys, 1989), since it has always been recognized that genetic variance becomes relatively greater as a consequence of decreasing environmental variance—which is where the second part of the claim comes in: High levels of proximal processes decrease environmental variance because they serve to supply interactive experiences to children who otherwise might not have them. This has the effect of not only increasing h^2 but also levelling group differences.

According to the bio-ecological view, it is also important to assess the dimensions of the child's ecology or environment because in two ways it provides limits on the efficiency of proximal processes. First, the environment contains the resources that need to be imported into proximal processes for the latter to work maximally. For example, it is not enough for parents to engage their adolescents in reciprocal interactions that sustain their children's attention while studying algebra (an example of the proximal process called "monitoring"), if the parents themselves cannot effectively explain the relevant algebraic concepts. Thus, the larger environment can sometimes set limits on proximal processes' efficacy.

A second reason for the importance of the larger environment is that it provides the stability necessary to benefit from proximal processes. A large literature (see Bronfenbrenner et al., in press) illustrates that regardless of social class, ethnicity, or ability levels, the less stable the environment, the worse the developmental outcome. Frequent changes in day care arrangements, stepparents, schools, or neighborhoods, for example, are associated with adverse outcomes, and this is presumably independent of the level of proximal processes. Measures of heritability are extremely sensitive to secular trends, generally dropping during times of economic scarcity and climbing during times of plenty (see Bronfenbrenner et al., in press). This is consistent with the view that the ecology brings to fruition differing levels of genetic potential, and h^2 s will fluctuate by a factor of three in conjunction with economic fluctuations. We assume that in times of economic scarcity the levels of proximal processes are reduced because caregivers' attention is deployed elsewhere.

The fourth important prong of the bio-ecological view is the incorporation of *motivation* as a key ingredient in its explanation of empirical findings. Briefly, an individual must not merely be endowed with some biological potential for a given cognitive resource, or merely be exposed to an environment

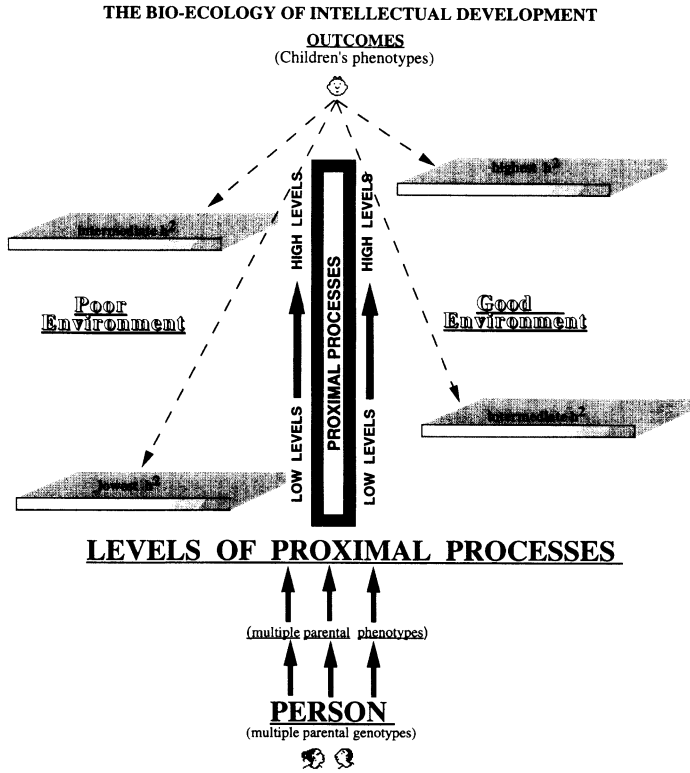


Figure 1. A partial depiction of the *Bio-ecological model*. At bottom are various genetic potentials, each of which is associated with a specific type of parental behavior (gene-environment correlation). Each genetic potential is actualized to the degree that proximal processes exist in sufficient form as well as resources from the more distal environment, thus leading to differential expectations about the magnitude of heritability. Missing from the figure are the individual pathways from each specific genetic potential.

that facilitates the expression of this cognitive resource; the individual must also be motivated to benefit from exposure to such an environment. Men in Ceci and Liker's (1986) study who demonstrated highly complex forms of reasoning at the racetrack did not exhibit the same degree of complex reasoning in other domains. Had they been exposed to environments that were conducive to, say, learning science or philosophy, and motivated to take advantage of such environments, they would have undoubtedly acquired the ability to think as complexly in those domains as they did at the racetrack, given the isomor-

phism that was built in between the type of reasoning needed to handicap a race and to reason scientifically in that experiment.

Figure 1 is a partial schematic of the bio-ecological trajectory and summarizes and elaborates some of the main points just made. As can be seen, the flow starts at the bottom with parental genes giving the child its early impetus and direction, ultimately influencing various forms of cognitive processing that are fairly independent of each other (depicted as three independent arrows emanating out of the child). The manifestations of these multiple gene-based resource pools are influenced by multiple forms of activities that caregivers engage in with children. One such activity, namely proximal processes, is postulated to actualize genetic potentials. Thus, variations in the level of resources in the environment (dichotomized for ease of illustration) along with variations in the level of proximal processes lead to different levels of h^2 .

There are several important aspects about this diagram. One is that proximal processes are more than expressions of parental genetics, otherwise there would be no basis for making differential predictions about the size of h^2 because all that we would need to know is the parental genotype. Yet, there are predictions about the differential size of h^2 s within the same levels of consanguinity when the levels of proximal processes differ (see later).

Second, the central arrow of this figure is broken rather than solid to emphasize a core principle underlying the bio-ecological model, namely that the influence of genetics and environment are never wholly separable. From the moment of conception, the actualization of inherited cognitive dispositions for embryological development do not occur in a vacuum, but are differentially responsive to the intrauterine environment (as well as to the intercellular environment, including interactions among hormones, inducers, enzymes, and proteins). This power of inherited propensities is not diminished after birth, because as the child interacts with persons, symbols, and objects in her environment, the latter becomes genetically loaded, as the active organism selects, changes, and hence constructs her own environment.

Third, this figure illustrates the stipulation that proximal processes are the engines of intellectual development, with higher levels of proximal processes associated with increasing levels of intellectual competence. Thus, these processes not only increase h^2 and reduce group differences, but they produce more competent organisms across the board.

Finally, the figure portrays some of the major predictions regarding the inclusion of proximal processes into a theory of intellectual development. Proximal processes are posited to exert a more potent influence on the various cognitive operations than does the larger environment (e.g., SES) in which they operate. Accordingly, the differences in intellectual outcomes (and their corresponding h^2 estimates) between poor and good environments are expected to be systematically smaller than the differences that are associated with low versus high levels of proximal processes. (This prediction is revealed in the compari-

son of the distances in the figure; distances between good and poor processes are greater than between good and poor environments.) In addition to the prediction that the highest magnitudes of h^2 are to be found under good-high conditions (i.e., high proximal processes–good environmental resources)¹, the bio-ecological model makes a corollary prediction; the largest *differences* in the magnitudes of h^2 for competent outcomes are to be found between children in the high-good condition and the low-poor condition. Hence, in the figure this is indicated by the largest physical distance between these two conditions.

Elsewhere, we have attempted to provide a limited test of these predictions with available data. For example, nearly 20 years ago, Bronfenbrenner (1975) reanalyzed the then available data on MZ twins reared apart. While there were no measures of proximal processes available, the resources in the environment probably differed markedly as a function of the ecologies in which the co-twins were raised. Bronfenbrenner reported that the intraclass correlations for IQ reached the high .80s when separated twins were reared in similar ecologies, but plummeted to .28 when they were reared in vastly different ecologies (e.g., agricultural/mining towns vs. manufacturing towns). These data are consistent with the bio-ecological model's prediction that h^2 will be both far higher and far lower than previously reported when the level of proximal processes and environmental resources are systematically varied.

An implication of the foregoing is that the greater the genotypic *dissimilarity*, the greater will be the impact of proximal processes in increasing phenotypic *dissimilarity*. That is, proximal processes will have their greatest impact in making those who are genetically dissimilar even more different, and in making those who are genetically similar even more similar, hence, increasing the value of h^2 . This is why the difference between the DZ and MZ intraclass correlations are almost always greater in good environments than in poor ones (see Bronfenbrenner et al., 1993). Thus, according to the bio-ecological view, the reason that h^2 is so high in studies of identical twins is not because the cognitive phenotypes of MZ pairs are unaffected by the environment, but because the cognitive phenotypes of the DZ pairs are so affected by the environment (e.g., Fischbein, 1982). Consequently, the bio-ecological model finds support for the interplay between ecology and biology in some of the very findings that have served as cornerstones of the behavior genetic and psychometric traditions.

The results of reanalyses with more recent data are also consistent with predictions from our bio-ecological model, though clearly much more needs to be done before this can be accepted as definitive (Bronfenbrenner et al., 1993). If

¹ Actually, this prediction is somewhat more complicated than stated here, and there are some circumstances where h^2 will be higher in poor environments (see Fischbein, 1982, for the specific area of arithmetic tests). But these are exceptions and we will refrain from delving into them here for ease of exposition.

we are correct about the critical role of proximal processes, then under differing levels of proximal processes, the size of the heritability estimate will change, possibly dramatically, because h^2 reflects only the proportion of "actualized" genetic potential, leaving unknown the amount of genetic potential that remains unactualized due to limited proximal processes.

Having described the basic features of the bio-ecological view of intellectual development, it now can be contrasted with the biological, environmental, and psychometric views of intelligence on three grounds. First, the bio-ecological view proposes multiple cognitive abilities, rather than one pervasive general factor, and these multiple abilities are at best only imperfectly gauged by tests of so-called general intelligence.

Second, the bio-ecological view differs from traditional nature–nurture interactionist views in the conception of the interaction between biological and environmental factors. The traditional interactionist view is that "genes encode phenotypes," and it is the eliciting power of the environment that releases the phenotypes. Thus the traditional interactionist view gives primacy to the importance of biological factors. The bio-ecological view departs from this preformationist view of the genotype and gives the environment a much more significant role. Genes do not encode phenotypes, but rather they manufacture proteins and enzymes that influence the expression of neighboring genes as well as interacting among themselves. If this view is correct (i.e., that the proteins and enzymes that are produced interact among themselves)—and there is every reason to believe that it is—this implies that such interactions are governed by physical and chemical laws independent of the strand of DNA from which they originated (see chapters in Subtelny & Green, 1982). A model of the translation of genotypes into phenotypes requires that we consider not just the proteins that genes manufacture but also the developmental role such proteins play, since most of the hormones, inducers, and inhibitors are connected in complex ways with the activity of multiple gene systems. Thus, the resultant morphology is only indirectly related to genes, making it impossible to explain phenotypes or morphological change exclusively or even primarily in terms of genes. As the evolutionary biologist Alberch (1983) wrote: "Even if we knew the complete DNA sequence of an organism, we could not reconstruct its morphology. We need to know about the epigenetic interactions that generated the phenotype" (p. 862).

Third, like other interactionist views of development, the bio-ecological view argues that the efficiency of cognitive processes depends on aspects of the context. However, according to Ceci (1990a), context is not an adjunct to cognition, but a constituent of it. Unlike traditional cognitive science, which has assumed that context is merely a background for cognition, the bio-ecological view regards context as an inextricable aspect of cognitive efficiency. Here context is defined broadly to include not only external features of the near and far environments and their motivational properties, but internal features of the

organism's mental representation, such as the manner in which a stimulus or problem is represented in memory. Thus, speed in recognizing letters and numerals depends on how those stimuli are represented in memory, with more elaborate representations leading to faster recognition rates (Ceci, 1990b). This explains why the same cognitive ability, no matter how basic, often operates inconsistently across diverse contexts (Ceci, 1990b). The same individuals who are slow at recognizing a stimulus in one domain may recognize it in another domain more quickly if its representation in the latter domain is more elaborate. In short, cognition-in-context research has shown that context, including the mental representation or mental context of a task, helps determine the efficiency of cognition.

Fourth, the bio-ecological view assumes that there are noncognitive abilities which are highly important for subsequent intellectual development, and which are inherited. For example, a child may inherit various types of temperament (e.g., restless, impulsive), physical traits (e.g., skin color, facial shape), and "instigative characteristics" (e.g., the "reward-seeking" type; Bronfenbrenner, 1989) that may influence later learning and development. While these traits are themselves influenced by gene systems, and can be shown to exert direct as well as indirect effects on subsequent IQ performance and school success, they are not cognitive in nature. So, these noncognitive characteristics and abilities can account for heritability patterns (e.g., IQs that run along consanguinity lines) without claiming that this is a consequence of the inheritance of a central nervous system with a determinate signal-to-noise ratio that limits processing capacity. Family members that share these characteristics may perform similarly as a result of their noncognitive dispositions, rather than because they share the same rate-limiting nervous systems (Lehrl & Fischer, 1988, 1990). Importantly, accounts of rate-limiting cognitive functioning that are based on EEG power-spectral density measures (e.g., Weiss, 1990a, 1990b), blood glucose levels in the brain, central nerve conduction velocity and oscillation (Jensen, *in press*), and heritability analyses (e.g., Pedersen et al., 1992), cannot distinguish between cognitive (i.e., inherited limits on CNS functioning) and noncognitive bases of performance. This is not to deny the importance of genes in intellectual performance, but merely a caution that just as it is the case that not all that is intellectual is genetic in origin, not all that is genetic is intellectual in nature.

Finally, the bio-ecological view departs from traditional behavior genetic models regarding the nature and meaning of h^2 . According to the bio-ecological view, h^2 reflects the proportion of "actualized" genetic potential, leaving unknown and unknowable the amount of unactualized genetic potential (Bronfenbrenner et al., *in press*). As we have argued above, nothing can be said about an individual's potential for success or failure without knowing about the level of proximal processes and more distal resources that exist in the child's environment. Proximal processes are the engines that drive development; they are the mechanisms that translate genes into phenotypes

(Bronfenbrenner et al., in press). If there are insufficient proximal processes in one's life, then h^2 will reflect only that portion of one's potential that can be brought to fruition by the limited level of proximal processes.

In sum, current research on intelligence and intellectual development points in a direction that emphasizes the role of context in the formation and assessment of an individual's manifold cognitive potentials. Although traditional measures of general intelligence possess good predictive validity in school, job, and training situations, this does not necessarily reflect the fact that a singular resource pool underpins a significant portion of the prediction, that the size of heritability estimates reflects the amount of variation due to purely genetic processes, or that only cognitive factors are involved. Thus, *prediction* and *explanation* can be fundamentally disjunctive enterprises in science. The bio-ecological view of intellectual development aims to fulfill the promissory note of the interactionist perspective, by taking cognizance of the biological and cognitive findings that have been reported throughout this century, but recasting them in a new developmental-contextual light.

REFERENCES

- Alberch, P. (1983). Mapping genes to phenotypes, or the rules that generate form. *Evolution*, 37, 861–863.
- Barrett, G. V., & Depinet, R. L. (1991). A reconsideration of testing for competence rather than for intelligence. *American Psychologist*, 46, 1012–1024.
- Bronfenbrenner, U. (1975). Nature with nurture: A reinterpretation of the evidence. In A. Montague (Ed.), *Race and IQ*. New York: Oxford University Press.
- Bronfenbrenner, U. (1989). Ecological systems theory. In R. Vasta (Ed.), *Annals of child development: Six theories of child development: Revised formulations and current issues* (pp. 185–246). Greenwich, CT: JAI Press.
- Bronfenbrenner, U., Ceci, S. J., & Lenzenweger, M. (1993). *Nature-nurture reconceptualized in developmental perspective: Toward a new theoretical and operational model*. Unpublished manuscript, Cornell University, Ithaca, NY.
- Bronfenbrenner, U., Lenzenweger, M., & Ceci, S. J. (in press). Heredity, environment, and the question "how"?: A first approximation. In R. Plomin & G. McClearn (Eds.), *Nature, nurture, and psychology*. Washington, DC: American Psychological Association.
- Carroll, J. B. (1992). Cognitive abilities: The state of the art. *Psychological Science*, 3, 266–270.
- Cavalli-Sforza & Bodmer. (1971) *The genetics of human populations*. San Francisco: Freeman.
- Ceci, S. J. (1990a). On the relationship between microlevel and macrolevel cognitive processes: Worries over current reductionism. *Intelligence*, 14, 1–19.
- Ceci, S. J. (1990b). *On intelligence...more or less: A bio-ecological treatise on intellectual development*. Englewood Cliffs, NJ: Prentice-Hall (Century Psychology Series).

- Ceci, S. J. (1991). How much does schooling influence cognitive and intellectual development? *Developmental Psychology*, *27*, 703–722.
- Ceci, S. J., & Liker, J. (1986). A day at the races: A study of IQ, expertise, and cognitive complexity. *Journal of Experimental Psychology: General*, *115*, 255–266.
- Ceci, S. J., & Ruiz, A. (1992). Cognitive complexity and generality: A case study. In R. Hoffman (Ed.), *The psychology of expertise: Cognitive research and empirical AI* (pp. 41–55). New York: Springer-Verlag.
- Ceci, S. J., & Ruiz, A. (1993a). Transfer, abstractness, and intelligence. In D. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 168–191). Norwood, NJ: Ablex.
- Deary, I. J., & Pagliari, H. C. (1991). The strength of *g* at different levels of ability: Have Detterman and Daniel rediscovered Spearman's "Law of Diminishing Returns"? *Intelligence*, *15*, 247–250.
- Detterman, D. K. (1986). Human intelligence is a complex system of separate processes. In R. J. Sternberg & D. Detterman (Eds.), *What is intelligence?* (pp. 57–61). Norwood, NJ: Ablex.
- Detterman, D. K. (1991). Reply to Dear and Pagliari: Is *g* intelligence or stupidity? *Intelligence*, *15*, 251–255.
- Detterman, D. K. (1993). Transfer on trial: The case for the prosecution. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 1–24). Norwood, NJ: Ablex.
- Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. *Intelligence*, *13*, 349–359.
- Detterman, D. K., Mayer, J. D., Caruso, D. R., Legree, P. J., Conners, F. A., & Taylor, R. (in press). The assessment of basic cognitive processes in relationship to cognitive deficits. *American Journal of Mental Retardation*.
- Detterman, D. K., & Persanyi, M. (1990, April). *Mental tests correlate highest in low IQ groups: Evidence from KABC*. Paper presented at the annual meeting of the American Educational Research Association, Boston, MA.
- Eysenck, H. J. (1988). The biological basis of intelligence. In S. Irvine & J. Berry (Eds.), *Human abilities in cultural context* (pp. 87–104). New York: Cambridge University Press.
- Fischbein, S. (1982). IQ and social class. *Intelligence*, *4*, 51–63.
- Gottfredson, L. S. (1986). Societal consequences of the *g* factor in employment. *Journal of Vocational Behavior*, *29*, 379–410.
- Hendrickson, A. E., & Hendrickson, D. E. (1982). The psychophysiology of intelligence. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 151–228). New York: Springer-Verlag.
- Herrnstein, R. J. (1973). *IQ and the meritocracy*. Boston, MA: Little, Brown.
- Horn, J. (1989). Models of intelligence. In R. Linn (Ed.), *Intelligence: Measurement, theory, and public policy* (pp. 29–73). Urbana, IL: University of Illinois Press.
- Humphreys, L. (1962). The organization of human abilities. *American Psychologist*, *17*, 475–483.
- Humphreys, L. (1979). The construct of general intelligence. *Intelligence*, *3*, 105–120.
- Humphreys, L. (1989). Commentary. In R. Linn (Ed.), *Intelligence: Measurement, theory, and public policy* (pp. 29–73). Urbana: University of Illinois Press.

- Hunter, J. (1983). *The dimensionality of the General Aptitude Test Battery and the dominance of general factors over specific factors in the prediction of job performance in the U.S. Employment Service* (Uses Test Res. Rep. #44). Washington, DC: U.S. Dept. of Labor, Employment, and Training Administration, Division of Counselling and Test Development.
- Hunter, J., Schmidt, F., & Rauschenberger, J. (1984). Methodological, statistical, and ethical issues in the study of bias in psychological tests. In C. R. Reynolds & R. T. Brown (Eds.), *Perspectives on bias in mental testing* (pp. 41–97). New York: Plenum.
- Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.
- Jensen, A. R. (1979). *g*: Outmoded theory or unconquered frontier? *Creative Science Technology*, 2, 16–29.
- Jensen, A. R. (1992). Commentary: Vehicles of *g*. *Psychological Science*, 3, 275–278.
- Lehrl, S., & Fischer, B. (1988). The basic parameters of human information processing: Their role in the determination of intelligence. *Personality and Individual Differences*, 9, 883–896.
- Lehrl, S., & Fischer, B. (1990). A basic information psychological parameter for the reconstruction of concepts of intelligence. *European Journal of Psychology*, 4, 259–286.
- Lickliter, R., & Hellewell, T. B. (1992). Contextual determinants of auditory learning in Bobwhite quail embryos and hatchlings. *Developmental Psychobiology*, 17, 17–31.
- Marshalek, B., Lohman, D., & Snow, R. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence*, 7, 107–127.
- Pedersen, N. L., Plomin, R., Nesselroade, J., & McClearn, G. E. (1992). A quantitative genetic analysis of cognitive abilities during the second half of the life span. *Psychological Science*, 3, 346–353.
- Plomin, R. (1985). Behavioral genetics. In D. Detterman (Ed.), *Current topics in human intelligence* (Vol. 1). Norwood, NJ: Ablex.
- Plomin, R., & Bergeman, C. S. (1991). The nature of nurture: Genetic influence on “environmental” measures. *Behavioral and Brain Sciences*, 14, 373–427.
- Plomin, R., & Neiderhiser, J. M. (1992). Genetics and experience. *Current Directions*, 1, 160–163.
- Scarr, S. (1989). Protecting general intelligence: Constructs and consequences for interventions. In R. L. Linn (Ed.), *Intelligence* (pp. 74–115). Chicago: University of Illinois Press.
- Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Development*, 63, 1–19.
- Schafer, E. W. P. (1987). Neural adaptability: A biological determinant of *g*-factor intelligence. *Behavioral & Brain Sciences*, 10, 240–241.
- Schneider, W., Körkel, J., & Weinert, F. E. (1989). Expert knowledge and general abilities and text processing. In W. Schneider & F. Weinert (Eds.), *Interactions among aptitudes, strategies, and knowledge in cognitive performance*. New York: Springer-Verlag.
- Siegler, R. S. (1989). Mechanisms of cognitive development. *Annual Reviews of Psychology*, 40, 353–379.

- Smith, G. J., & Spear, N. E. (1978). Effects of home environment on withholding behaviors and conditioning in infant and neonatal rats. *Science*, *202*, 327–329.
- Spearman, C. (1904). General intelligence objectively determined and measured. *American Journal of Psychology*, *15*, 206–221.
- Sternberg, R. J. (1985). *Beyond IQ: The triarchic theory of intelligence*. New York: Cambridge University Press.
- Sternberg, R. J. (1990). *Metaphors of mind*. New York: Cambridge University Press.
- Subtelny, S., & Green, P. B. (1982). *Developmental order: Its origin and regulation*. New York: Alan R. Liss.
- Walker, C. H. (1987). Relative importance of domain knowledge and overall aptitude on acquisition of domain-related information. *Cognition and Instruction*, *4*, 25–42.
- Weiss, V. (1990a). From short term memory capacity toward the EEG resonance code. *Personality and Individual Differences*, *10*, 501–508.
- Weiss, V. (1990b). The spatial metric of brain underlying the temporal metric of EEG and thought. *Gegenbauers Morphology Jahrb. Leipzig*, *136*, 79–87.

Chapter 5

***Theoretical Possibilities:
The Relation of
Human Intelligence to
Basic Cognitive Abilities***

Douglas K. Detterman

Case Western Reserve University

Human intelligence has been studied for over 100 years, yet seems as elusive a concept now as when first studied. One reason for this elusiveness is that the relationship of basic cognitive abilities to complex tests of human intelligence is poorly understood. The purpose of this chapter is to enumerate the theoretical possibilities for the relationships between standard measures of intelligence (IQ tests) and basic cognitive abilities and to describe the sorts of evidence that can be used to choose between the possible positions.

EARLY HISTORY OF THE STUDY OF BASIC MENTAL ABILITIES

The relationship between basic cognitive abilities and more complex measures of intellectual functioning has always been of concern to those interested in intelligence. Galton (1883) was one of the first to recognize a relationship between basic cognitive abilities and intelligence. Between 1884 and 1890, Galton operated a laboratory in the South Kensington Museum (now the Victoria and Albert Museum, though not in the present building of this museum). For a fee, visitors to the museum could have their basic anthropometric measurements recorded. Besides providing a repository for family information, Galton's idea was to determine how very basic measures like sensory acuity and reaction time related to accomplishment. Though Galton never analyzed the data, Johnson et al. (1985) did and found them to be surprisingly consistent with data collected recently using much more sophisticated equipment and methods. There is a tendency to regard Galton's efforts as a historical curiosity, notable because data were collected in a museum and because subjects paid to participate. Even though this work is historically interesting, it is also scientifically important. Galton did what scientists are doing now well before it became fashionable. Current interest in research on the cognitive basis of individual differences in intelligence did not begin until the 1960s.

Early Studies Inhibit Search

If the relationship between basic abilities and intelligence was of such concern to researchers, why was the issue abandoned and not taken up again until recently? Tuddenham (1963) provides a compelling explanation of why academicians abandoned the issue. McKeen Cattell announced his intention to study individual differences in basic abilities and even published a listing of tests to be used in this effort (Cattell, 1890). The tests included bisection of lines, auditory reaction time, least perceptible difference for weights, two-point threshold, and color naming. Cattell acknowledges his debt to Galton and gives his purpose for applying these tests to University of Pennsylvania students in no uncertain terms:

Psychology cannot attain the certainty and exactness of the physical sciences unless it rests on a foundation of experiment and measurement. A step in this direction could be made by applying a series of mental tests and measurements to a large number of individuals. (p. 373)

Cattell's proposal evidently generated a great deal of enthusiasm, because by 1910 there was a sufficiently large number of basic mental tests to publish a

manual filled with a wide variety of cognitive tests (Whipple, 1910). Only four years later, this manual had to be expanded to two volumes. But Cattell had already planted one of the seeds that would destroy this early interest in the relationship of basic cognitive abilities to more complex mental skills.

When Cattell moved to Columbia, he administered his modified test battery to an entire incoming freshman class. It was his intention to correlate these basic measures with grades and graduation standing four years later. Cattell was so confident of the results that he discussed the effect they would have before they were analyzed. Columbia hosted summer seminars for educators from around the country and Cattell lectured in these seminars. He suggested that the data he were collecting would provide a model for a “science” of education.

When the data finally were collected, the low correlations came as a shock to Cattell. He had expected much higher correlations between the individual cognitive tasks and school performance. Disappointed, he dissociated himself from the data. Wissler (1901), a graduate student working with Cattell, was given the data to analyze and published the results. Cattell’s name did not appear on the publication. Wissler was somewhat more optimistic than Cattell about the results:

We have found low degrees of correlation in many cases, which seems to imply that, if functional relations hold in these tests, they are exceedingly complex, even more so than many psychologists have assumed. This promises little for such tests from a practical point of view. While the tests do seem to have some value when applied to children in the lower schools, they tell us nothing as to the general individual worth of college students or of adults. Indeed, they lead us to doubt the existence of such a thing as general ability. This general negative statement is likely to impress the reader in such a way that he will feel disposed to declare that psychological tests are of no value, and that time spent in making them is mere waste. The writer cannot share in such a feeling. (pp. 54–55)

Wissler went on to explain how devising mental tests would suggest hypotheses to clarify the nature of individual differences.

Unfortunately, the Wissler report followed on the heels of another influential negative finding published three years earlier. Sharp (1898) was a student in Titchener’s laboratory and she was interested in mental tests. She tested seven fellow graduate students on a number of tests of memory, imagery, discrimination, and attention. Her results indicated no relationship between test results and academic performance. The graduate students obviously represented a highly curtailed range of intellectual ability, but no one seems to have recognized that this methodological problem would have substantially reduced the size of the correlations obtained. It is not surprising that Sharp did not realize the problem because correlation was a new technique. It is surprising that

many investigators today still make the mistake of using samples with curtailed ranges even though the effects of restricted ranges on correlation coefficients are well known.

According to Tuddenham, the Sharp and Wissler studies made the study of mental tests an unpopular undertaking in academic circles, particularly when the tests were of basic mental processes. But there were also other pressures that led investigators away from studying the basic processes that underlie more complex mental performance. One of the most important of these was that complex tests seem to predict important everyday criteria like school achievement better than basic tests.

Binet and Henri (1896) had proposed a program of research to assess ten faculties or basic abilities, including memory, attention, comprehension, and imagination. In the first complete intelligence test developed by Binet and Simon (1908), the influence of the earlier theory can be seen in the choice of some of the items: comparing two lines of equal length, comparing two weights, memory for pictures, placing five weights in order, digit span, and drawing a design from memory. In the evolution of the test, most of these basic items were replaced with more complex ones simply because more complex items were better at predicting the criterion of school achievement. The fact that complex items work better than less complex items led researchers away from the basic mental abilities investigated by Galton, Cattell, and other early investigators.

The Reawakening of Interest in Basic Cognitive Abilities as Explanations of Complex Mental Processes

In the last two or three decades, the word *mental* has become much more fashionable than it was at any time during the preceding 50 years. With interest in higher mental processes has come a renewed interest in the relationship between basic cognitive processes and higher mental processes. Interest has come from several directions and for different reasons but each line of research adds to the importance of understanding the relationship between basic cognitive abilities and higher mental processes.

One line of research where the question of the relation between basic cognitive abilities and intelligence was of importance was in mental retardation research. Investigators like Zeaman and House (1963) and Ellis (1963) suggested that the poor performance of mentally retarded persons was due to deficits in one or a small number of basic processes. The Ellis (1970) rehearsal deficit hypothesis has been one of the most extensively investigated. One problem with the deficit approach is that it appears that all, not just one or a few, processes are deficient in the mentally retarded. In reviewing studies of information processing in the mentally retarded, Detterman (1979) found that every

process which had been investigated showed a deficit in mentally retarded subjects when they were compared to nonretarded performance on the same task. The term *everything deficit* was suggested to summarize these findings. One obvious explanation for the everything deficit is that the models used to study the mentally retarded were not explicitly developed for the study of individual differences. It may be, that with more exact specification of basic abilities, deficits would be more precisely identified.

A second line of research which has reawakened interest in basic abilities is the cognitive components research of Sternberg (e.g., Sternberg & Spear, 1985) and the cognitive correlates approach of Hunt (e.g., Hunt, Lunneborg, & Lewis, 1974). The cognitive components approach was used to understand the basic processes that compose more complex mental tasks. Performance on intelligence test items was decomposed to its basic components. The cognitive correlates approach, on the other hand, sought to explain performance on a standardized test by correlating it with more basic cognitive tasks. Basic cognitive tasks were selected because they were thought to be involved in the more complex test. These two approaches have made the study of individual differences once again acceptable to cognitive and experimental psychologists who, until recently, had a low opinion of this kind of research.

Both the cognitive correlates and cognitive components approach have encountered similar difficulties. The basic processes identified using these approaches are correlated with each other. The standard explanation of this often substantial correlation is that the common variance among tasks represents meta- or higher order processes like strategy planning or self-monitoring (Detterman, 1984).

Another line of research that has focused attention on the relationship between basic cognitive abilities and higher mental processes is the work of Jensen (1979, 1980, 1982a; Jensen & Munro, 1979; Jensen, Schafer, & Crinella, 1981). Because standardized tests of intelligence depend on complex items, they are open to charges of bias. Complex items, at least on the face of it, do appear more subject to cultural influences than more basic tests of cognitive abilities. For this and other reasons, Jensen and others (see Carroll, 1980; Vernon, 1987) have looked for tasks of basic processes that were more basic and less subject to the objections leveled at standardized IQ tests. Reaction time has been the most extensively studied task by this group of researchers. The hope of some (see Eysenck, 1982) is that these basic tasks will show that a single variable, such as mental speed, underlies performance on basic cognitive tasks and more complex tasks, too.

All researchers attempting to understand individual differences have made implicit assumptions about the relationship of basic abilities to higher mental processes. Despite the extensive interest in the relationship of basic cognitive abilities to more complex mental tests, there have been few attempts to consider alternative theoretical positions in a general context. Most of the remain-

der of this chapter is devoted to exactly that question: How do basic cognitive processes predict and explain more complex measures of mental functioning, such as those represented in IQ tests? What theoretical alternatives are available and what evidence discriminates among them? Before considering theoretical alternatives, however, a clear statement of exactly what is to be predicted is needed.

DEFINITION OF GENERAL INTELLIGENCE

Spearman (1904, 1927; Hart & Spearman, 1912) noted that when a battery of mental tests was given, all of the tests in the battery were positively intercorrelated. He referred to this correlation as positive manifold and thought that it might represent a single underlying variable. Figure 1 shows how Spearman

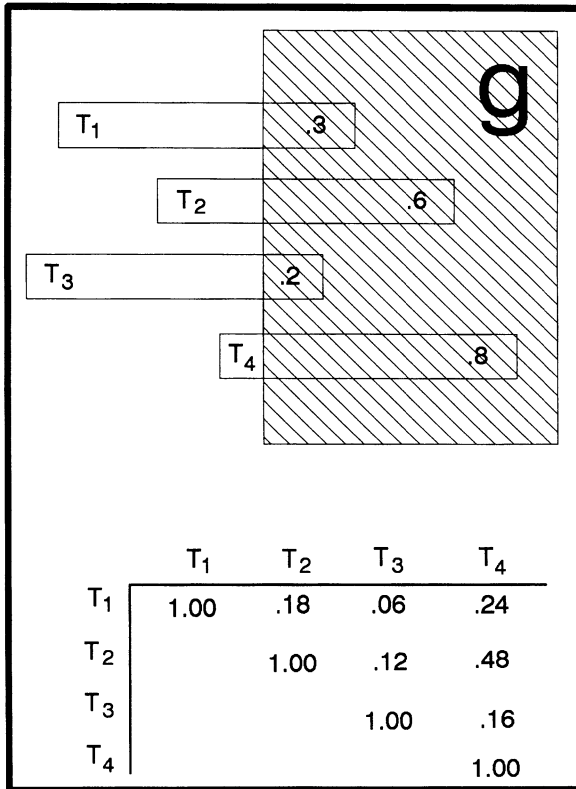


Figure 1. A schematic representation of the correlation of four tests with a general factor and the correlation matrix that would result from them.

conceptualized this model. He assumed a general, underlying factor correlated more or less with all mental tests. He called this factor g for general intelligence. The degree of correlation of each test with g is shown by the extension of the bar beyond the line. If tests shared variance then they should correlate with each other. The size of the correlation between the two tests would depend on how much of common, g variance they shared. In other words, if tests measure the same underlying factor, their correlation is a function of the extent each test measures this factor.

The correlation matrix shown in Figure 1 is constructed on Spearman's theoretical principle. Test 1 correlates .3 with the underlying factor and Test 2 correlates .6 with the same factor. The correlation between them will be .18 (.3 x .6). The correlation of every pair of tests can be computed as a product of the two tests' correlations with the underlying factor. Note that the value of the correlation of each test with the underlying factor is its correlation with g , or general intelligence. (Today this correlation is called a factor loading.)

Spearman needed to show that the correlations between all tests in a matrix were a function of a common underlying factor. He reasoned that the correlation of two tests was the product of their correlation with the underlying g factor. That is, the correlation of Test 1 and Test 2 could be represented by the product of Value 1 and Value 2, where each value is the test's correlation with the general factor. Products of test correlations could also be represented by a series of products. For example, the product of the correlation of Test 1 and Test 2 (r_{12}) and the correlation of Test 3 and Test 4 (r_{34}) could be represented as:

$$r_{12} * r_{34} = (V_{-1} * V_{-2}) * (V_{-3} * V_{-4}) \quad (1)$$

The subscripted Vs refer to the value of g (the factor loading) for each test in Figure 1. V_{-1} is .3 for T-1. The product of r_{13} and r_{24} would be:

$$r_{13} * r_{24} = (V_{-1} * V_{-3}) * (V_{-2} * V_{-4}) \quad (2)$$

Since the order of multiplication is unimportant, the quantities expressed in Equation 1 and 2 are equal. Both Equations 1 and 2 include all the values V_{-1} , V_{-2} , V_{-3} , and V_{-4} on the right-hand side. Therefore, Spearman reasoned, the quantity on the right-hand side of Equation 1 is equal to the quantity on the right-hand side of Equation 2 because they are both products of the same terms. If Spearman's assumptions about the right-hand side of each equation were correct, then the left-hand side of Equation 1 should equal the left-hand side of Equation 2. If the left-hand sides are equal to each other, then the difference between them should be zero:

$$r_{12} * r_{34} - r_{13} * r_{24} = 0 \quad (3)$$

Spearman called this a tetrad difference. If a matrix of test correlations was completely determined by a single underlying factor, then all of the tetrad differences would be zero. If there was not a single factor as shown in Figure 1 but many separate factors, then the differences would not equal zero.

Spearman's method of tetrad differences was the forerunner of factor analysis. It was the first method to demonstrate the existence of the general factor or g . The same reasoning that Spearman used can be applied to more complex problems where the correlations of the tests are determined by more than one underlying factor. All of modern factor analysis is based on Spearman's fundamental insight. (Who, exactly, invented factor analysis is the subject of historical debate. People other than Spearman had similar insights at about the same time.)

More complex and sophisticated methods of factor analysis have been developed since Spearman's method of tetrad differences. Today the most common definition of the general factor is the first principal factor (Jensen & Weng, in press). Despite increased sophistication, the concept of g has not changed since Spearman proposed it. It is a measure of the degree of correlation among all of the tests of a battery with a common factor.

The concept of general intelligence, as conceptualized by Spearman, is a purely statistical concept. Though Spearman felt that the common variance shared among tests in a battery resulted from a single cause which he called mental energy, there are certainly many other potential explanations for this statistical finding.

EVIDENCE FOR IMPORTANCE OF GENERAL INTELLIGENCE OR g

Though g is a statistical concept, that doesn't make it any less real. The evidence for the general factor, g , is overwhelming. When a battery of standardized intelligence tests is factor analyzed, the first principal component typically accounts for 40% to 80% of the total variance. The general factor is so pervasive and so large that nearly every factor analytic theory of intelligence (with a few major exceptions) includes a general factor in one form or another.

The general factor also seems to carry most of the variance that makes IQ tests predictive of important validity criteria. For example, Jensen found that general factor scores correlated .38 with service school performance. This correlation went to only .41 when the entire battery, not just g , was used to predict performance. The general factor also seems to be the major source of differences between various cultural groups (Jensen, 1984). The more g -loaded a test is, the larger the differences between racial groups on the test will be.

Any acceptable theory of intelligence will have to account for g . However, there are purely empirical reasons to use a measure of g instead of total test

score as a criterion measure in the investigation of basic cognitive abilities. All tests and subtests contain sources of variance unique to their format or administration. Using g reduces the variance specific to a single test or subtest. If two tests given in the same manner correlate, the correlation between them could result from the similarity in method of administration, not because the two tests depend on the same mental processes. This is less likely when g is used because sources of variance unique to a particular test are not included in g . Using g can act like a filter, removing unique sources of variation present in a set of tests.

FACTS TO BE ACCOUNTED FOR BY RELATIONSHIP OF BASIC COGNITIVE ABILITIES TO GENERAL INTELLIGENCE

A major goal of any theory of intelligence is to find the basic process or processes of general intelligence, but there are a number of other findings such a theory should account for in addition.

Test Complexity and General Intelligence

More complex tests demonstrate larger differences in mental ability than simple tests. This is true even for rats. Lashley (1929) used two different mazes which differed in complexity in his classic ablation experiments. The effects of brain damage were much more obvious on the complex maze than on the simple one.

Jensen (1982b) compared performance on reaction-time tasks which differed in complexity. A simple item required the subject to move to a lighted button. In a complex item, subjects had to pick a correct synonym for a word and press the button under the word. In general, the more complex the task, the longer the reaction time and the higher the correlation with standard psychometric measures of intelligence.

As was pointed out earlier, intelligence tests have evolved from simple to more complex items. Binet included some very simple test items which were replaced in later tests with more complex items. An examination of intelligence tests currently in use supports this contention. Nearly every item on these tests is a complex one, even by the most conservative definition of complexity.

The fact that complex items make a "better" IQ test than simple items has led some to conclude that intelligence is related to complex reasoning. Some authors (Sternberg & Salter, 1982) have even gone so far as to suggest that basic cognitive abilities will be found to be unrelated to intelligence because they measure processes which are too simple. If intelligence is complex reasoning, then simple processes will not reflect it, the reasoning goes.

Low Relationships of Basic Abilities to IQ

Basic measures of cognitive abilities like measures of memory, attention, and learning have low correlations with standardized measures of intelligence. Typically, the correlations are less than .35. In fact, Sternberg (Sternberg & Salter, 1982) indicated the difficulty of obtaining high correlations between basic cognitive abilities and intelligence. Sternberg and Salter referred to this problem as the .30 barrier because correlations higher than this have seldom been obtained.

Although some investigators have obtained higher correlations between basic measures of cognitive functioning, these tend to be isolated, unreplicated findings. Any theory accounting for the relationship of basic ability and standardized measures of intelligence will have to explain this low relationship. As with the findings concerning complexity, some investigators have taken the low correlations to be an indication that basic cognitive abilities are unrelated to intelligence.

Unique Conditions

A number of unique conditions exist that should be explained by any theory relating basic abilities to more standardized measures. One class of conditions to be explained are those in which a particular ability is deficient or exceptionally outstanding. Mnemonicists, idiot savants, and persons with specific learning disabilities or focal brain damage would fall into this category. Exactly why do these phenomena occur? How can they be explained on the basis of individual differences in basic cognitive abilities?

Psychometric Theories Postulating Separate Factors

In the following section of the chapter, the basic theoretical possibilities are presented and formally developed. Although the development is somewhat tedious it is necessary to have an explicit statement of each position.

EXPLANATIONS OF GENERAL INTELLIGENCE IN TERMS OF BASIC ABILITIES

There are three possible models relating basic cognitive abilities to g . The three models are differentiated by the number of variables postulated to be necessary to explain g . For convenience, these models will be referred to as Models I through III. In Model I, only one basic cognitive variable is postulated to be necessary to explain g . In Model II, a small-to-moderate number of variables is

needed to explain g , and in Model III, a very large number of variables is required to explain general intelligence.

The following sections of the chapter provide an overview of each of the general theoretical positions. The initial discussion of each model is general. In later sections, a mathematical statement of this verbal description is provided along with derivations. After detailing all of the models, the various sources of evidence to be explained by these models will be presented briefly.

Model I

Model I is the model originally suggested by Spearman to account for the large correlation between tests. Spearman suggested that the important variable was mental energy, but he was never very specific about what that was. More recent theorists have been much more specific. Eysenck (1982) has proposed that the single variable accounting for intelligence is mental speed. Evidence supporting this claim includes high correlations between intelligence and some reaction-time tasks and high correlations between measures of average evoked potentials and intelligence.

Figure 2 is a diagrammatic representation of Model I. The large circle represents the common variance from a battery of standardized intelligence tests

Model 1

g as a Single Thing

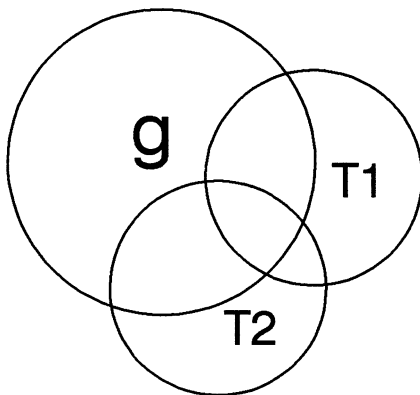


Figure 2. Model I, g as single thing. T_i represent basic cognitive tasks.

or, in other words, g . The smaller circles represent basic cognitive tasks correlated with g . For each of the models, there are two characteristics of particular interest. The first is the correlation between g , or the standardized tests (T), and the basic cognitive tasks (t). The second is the correlation between the basic cognitive tasks themselves.

Whatever the single variable is that causes g according to this model, it is the amount of this variable that the test or task measures that determines the correlations. A standardized test correlates with g to the extent it is g -loaded. This statement is circular, save for the fact that a given test correlates with g about the same over changes in the composition of the battery (Jensen & Weng in press). If Test 1 (T_{-1}) is included with tests T_{-2} to T_{-6} and then with tests T_{-7} to T_{-11} in principal factor analysis, the factor loading (which is the tests correlation with g) for Test 1 should be about the same in both batteries. Large changes in the factor loadings of T_{-1} across batteries would suggest that the definition of g changed with the composition of the battery. This change would be hard to explain if g was a single variable.

In Model 1, a cognitive task correlates with a standardized test (or g) for the same reason a standardized test correlates with g : it is more or less g -loaded. By g -loaded is meant the degree to which the basic cognitive task measures the single variable thought to produce g . For example, if mental speed is the variable causing g , cognitive tasks will have higher or lower correlations with g depending on the purity of the task in measuring this variable. A task more dependent on mental speed should be more g -loaded than tasks not so dependent on mental speed.

Standardized tests of intelligence are typically found to be more g -loaded than measures obtained from basic cognitive tasks. The goal of investigators adopting Model I classes of models is to find a basic cognitive task which is as pure a measure of the variable constituting g as it is possible to get. The rationale is that a basic cognitive task will more clearly demonstrate the nature of the underlying variable producing g . Therefore, the research strategy is to find basic cognitive tasks which correlate as highly as possible with g or standardized tests. Because these basic cognitive tasks are simpler, more direct representations of the underlying variable causing g , they constitute an explanation of general intelligence or, at the very least, a more reasonable possibility for finding an explanation than complex, standardized tests offer.

Basic cognitive tasks correlate with each other to the extent that they are g -loaded or that they measure some other common source of variance which is not g . The certain source of correlation between two basic cognitive tasks, though, is the amount of g -loadedness of each. If two basic cognitive tasks correlate with g , they must correlate with each other. This is an obvious conclusion because both tasks measure the same underlying variable if they correlate with g , and so must correlate with each other. The exact magnitude of the cor-

Model 2

g as a Several Things

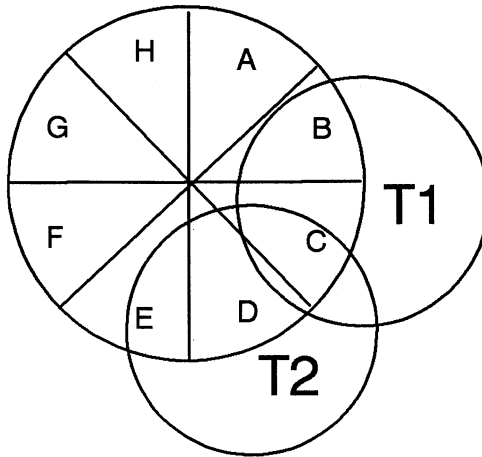


Figure 3. Model II, g as a small number of independent processes. T_1 and T_2 represent basic cognitive tasks.

relation between basic cognitive tasks can be estimated and is an important consideration in discriminating this model from others. This issue is considered in greater detail in a later section.

Model II

Model II is shown in Figure 3. Standardized tests of intelligence (and g) are composed of a small number of independent basic abilities. These abilities are represented by sections of the largest circle. Basic cognitive tasks are represented by smaller circles and measure some of these basic processes.

The correlations between more complex, standardized tests of intelligence are high because these tests measure all of the processes constituting g . A definition of g derived from this model is that g is all basic processes common to intellectual tasks. Because complex items are more likely to contain more processes, it is not surprising that the evolution of standardized intelligence tests has been toward complex items. While it may not be as clear what com-

plex items measure, they measure more of what is important in predicting criteria than simple items do.

The correlation of a basic cognitive task with g is dependent on the number of basic processes measured by the basic task or, more correctly, the number of processes included in the measure obtained from the basic cognitive task. The more processes the basic cognitive task contains, the higher its correlation will be with g .

In the same manner, basic cognitive tasks correlate with each other to the degree that they measure common processes or sources of common variance not included in g . In Figure 3, t_1 measures B, C, and D, and t_2 measures C, D, and E. These two tasks correlate because they measure common processes C and D. Note that there is no necessary relationship between the degree to which a basic cognitive task correlates with g and the size of its correlation with other basic cognitive tasks. It is entirely possible to have basic cognitive tasks which correlate perfectly with other basic cognitive tasks. They would be tasks which measure all of the same processes. Or it is possible to have basic cognitive tasks that correlate as highly with g as the perfectly correlated basic cognitive tasks and yet are uncorrelated with each other. These would be tasks that measure different processes. For example, if t_1 measured A, B, and C, and t_2 measured D, E, and F, both would correlate equally with g (assuming A to F accounted for the same amount of variance) but would not be correlated with each other.

Model II has, to my knowledge, never been formulated in the manner presented here. Thurstone's (1935, 1938) attempt to derive Primary Mental Abilities and Guilford's (1967) factor analytic model are different in a number of respects. Most importantly, though, both Thurstone and Guilford attempted to account for intelligence by using tests of about the same degree of complexity as standardized tests. According to this position, it would not be possible to account for complex tests with other complex tests. Further, neither Guilford's nor Thurstone's position was an attempt to account for g using more basic measures of cognitive functioning. Neither of these investigators took g into account in formulating their position and this proved a serious problem for both positions.

In Thurstone's case there were correlations between his mental ability factors which reflected g . Guilford's position, on the other hand, has been criticized because it is difficult to test (Undheim & Horn, 1977) and it is, therefore, nearly impossible to differentiate it from other theories including g -based positions.

Model III

Model III owes its initial existence to Thomson (1916, 1919, 1927, 1935, 1939), who originally formulated it to show that Spearman's g could be

accounted for by hypothesizing more than one underlying variable. After spending time with Thorndike (1927), who accounted for intelligence using associationism, Thomson referred to his theory as bond theory. According to Thomson's bond theory, g was accounted for by a large—approaching infinite—number of underlying variables representing number or strength of associations. A current proponent of a similar position is Humphreys (1979).

Figure 4 shows a representation of this position. The small dots represent bonds or associations. A test is regarded as a sample of these bonds. Two tests are correlated to the extent that they sample the same bonds. (Because tests are regarded as samples of bonds and because Thomson first simulated test results by tossing dice, Thomson's theory is sometimes called sampling theory.)

The extent to which two basic tasks are correlated is, likewise, a function of the number of common elements that they share. Since tests and basic cognitive tasks are assumed to be random selections of all bonds, the degree to which basic cognitive tasks are intercorrelated with each other is dependent on their correlation with g . Because a measure of g should be the best representa-

Model 3

g as a Large Number of Things

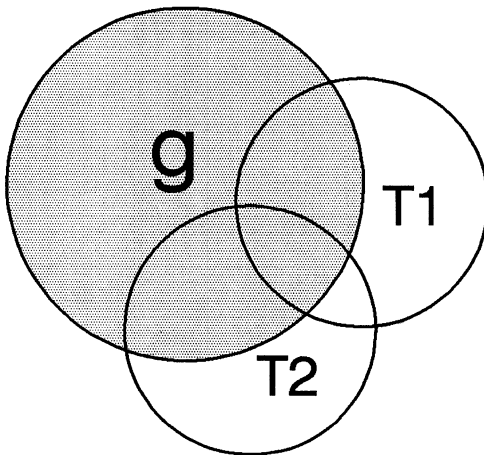


Figure 4. Model III, g as a very large number of processes. T_i represent basic cognitive tasks.

tion of the universe of bonds, the correlation of a basic cognitive task with g should be an indication of the number of bonds sampled by the task. By the laws of probability two samples of these bonds (i.e., two basic cognitive tasks) should overlap to an extent dependent on the proportion of bonds each samples. Since the number of bonds each basic cognitive task samples is its correlation with g , the correlation of two basic cognitive tasks is dependent on each task's correlation with g .

Similarity of Model I and III

Superficially, Model III bears a similarity to Model II because both models include more than one explanatory variable. But in fact Model III is really identical to Model I when the number of variables is very large to infinite. This was demonstrated by the work of Garnett (1919a, 1919b, 1920; Mackie, 1928; Maxwell, 1972) which seems to have been all but forgotten. Garnett showed that when the number of bonds (or whatever the underlying processes were thought to be) was large, Thomson's position was exactly equivalent to Spearman's.

Although Garnett's demonstration is quite intricate, it is possible to come to the same conclusion intuitively. Thomson's theory concerns drawing a sample from a large, nearly infinite population. As the sample from that population becomes large, the parameter being estimated (strength or number of bonds) will become less variable, approaching a fixed constant; g is also a similar estimate of an average value. Both theories postulate a single underlying variable. In Thomson's theory that variable is the average strength or number of bonds available to the individual as sampled by the test. Without reformulation, Thomson's theory is mathematically indistinguishable from Model I. Both estimate a single characteristic of the system. For that reason, Model I and Model III will be considered together in subsequent portions of this chapter.

Thomson's theory could most probably be differentiated from a single variable explanation of g by other means than those considered here. But if Model I cannot be supported with the evidence presented here, then neither can Model III. Discriminating between them would be a wasted effort. The most important issue is the discrimination of Models I and III from Model II.

DISCRIMINATION OF MODEL I AND III VS. MODEL II

In the preceding discussion, there was one obvious difference between the positions: In Model II, the correlations among basic cognitive tasks are not dependent on the tasks' correlation with g . For both Models I and III the correlations between basic cognitive tasks are directly related to each task's corre-

lation with g . If it were possible to determine the exact relationship between the correlation of g (or standardized tests as a representative of g) with basic cognitive tasks and the correlation of basic cognitive tasks, this information might be used to test the alternative positions.

Intertask Correlation

The following sections develop the statistical rationale for the relationship between r_{Tt1} and r_{t1t2} .

Model I and III. More explicit definitions are needed than have been used to this point. A standardized test as a representation of g is defined as follows:

$$T_i = \sqrt{P_G} * g_i + \sqrt{P_E} * E_i \tag{4}$$

T_i is the total test score of subject i . This score contains two components: g_i , the amount of g the subject has (expressed as a z-score to simplify later proofs) and E_i , an error score for the subject which is uncorrelated, across subjects, with g_i . For Model III, g_i represents the average number or strength of bonds. The weights, P_G and P_E , represent the amount of g and E measured by the test and sum to 1.0. The square roots of the weights are used to simplify the results. These weights are set for a particular test and do not change across subjects.

Basic cognitive tasks are defined as follows:

$$t_{i1} = \sqrt{p_{g1}} * g_i + \sqrt{p_{e1}} * e_{i1} + \sqrt{p_{o1}} * o_{i1} \tag{5}$$

$$t_{i2} = \sqrt{p_{g2}} * g_i + \sqrt{p_{e2}} * e_{i2} + \sqrt{p_{o2}} * o_{i2} \tag{6}$$

Subject i 's scores on two basic cognitive tasks are defined for t_{i1} and t_{i2} . Nearly all of the terms have the same meaning as in the definition of a standardized test except they are in lowercase to distinguish them from the standardized test. The terms, o_{i1} and o_{i2} , represent other sources of variance not in g , but which could cause a correlation between the basic cognitive tasks. The term g_i is exactly the same as that in the standardized test above so the correlation between the two terms is 1.0. What distinguishes basic cognitive tasks from each other, according to Models I and III, is the weighting of g . Some tasks are said to be more g -loaded than others and that is represented by the weights, p_{g1} and p_{g2} .

Given these definitions, it should now be possible to determine the correlation between a standardized test, T_i , and a basic cognitive task, t_{i1} .

$$r_{Tt1} = \sum(X_T * x_{t1}) / (N S_T S_{t1}) \tag{7}$$

Equation 7 is the deviation score formula for the correlation coefficient. Substituting the previously given definitions results in the following (because $r_{ge} = r_{go} = r_{eo} = 0$):

$$= \sum(\sqrt{P-G} * g_i * \sqrt{p-g1} * g_i) / (N S_T S_{t1}) \quad (8)$$

and because $S_T = S_{t1} = 1$ and because weights are constants:

$$= \sqrt{P-G} * \sqrt{p-g1} * (\sum(g_i)^2 / N) \quad (9)$$

Because g_i is in z-score form and the sum of the z-scores squared is equal to N , then:

$$r_{Tt1} = \sqrt{P-G} * p_{g1} \quad (10)$$

Not surprisingly the correlation between a standardized test and a basic cognitive task is a direct function of the weighting of g in each.

By similar development:

$$r_{t1t2} = \sqrt{p-g1} * p_{g2} \text{ if } p_{o1} = p_{o2} = 0 \quad (11)$$

Also not surprisingly, basic cognitive tasks are correlated to the extent each is weighted with g .

Using this information, it is possible to determine a lower limit for the correlation of basic cognitive tasks according to these two models. It was previously shown that:

$$r_{Tt1} = \sqrt{P-G} * \sqrt{p-g1} \quad (12)$$

or

$$\sqrt{p-g1} = r_{Tt1} / \sqrt{P-G} \quad (13)$$

The same would apply to a second basic cognitive task:

$$\sqrt{p-g2} = r_{Tt2} / \sqrt{P-G} \quad (14)$$

It was also previously shown that:

$$r_{t1t2} = \sqrt{p-g1} * \sqrt{p-g2} \quad (15)$$

By substituting the right-hand portion of Equations 12 and 13 for $\sqrt{p_{g1}}$ and $\sqrt{p_{g2}}$, the following is obtained:

$$r_{-t1t2} = (r_{-Tt1} * r_{-Tt2}) / (\sqrt{P_{-G}} * \sqrt{P_{-G}}) \tag{16}$$

or

$$= (r_{-Tt1} * r_{-Tt2}) / P_{-G} \tag{17}$$

If P_{-G} is 1.00 (a perfectly pure, reliable measure of g) then the correlation between two basic cognitive tasks is the product of each tasks correlation with g . This makes perfect intuitive sense. If P_{-G} is less than 1, then:

$$r_{-t1t2} \leq r_{-Tt1} * r_{-Tt2} \tag{18}$$

and so a lower limit for the correlation between basic cognitive tasks will be the product of their correlations with g (or, in practice, with a standardized test representing g).

Model II. Model II is more complicated. Assume that g consists of independent abilities A through H (excepting E which represents error) and that each subject has a certain amount of each of these abilities. Further assume that a standardized test measures all of these abilities in weighted amounts indicated by P (all Ps sum to 1.0). Then subject i 's total score, T_{-i} , could be represented as:

$$T_{-i} = \sqrt{P_{-A}} * A_{-i} + \sqrt{P_{-B}} * B_{-i} \dots \sqrt{P_{-H}} * H_{-i} + P_{-E} * E_{-i} \tag{19}$$

In a similar fashion, two cognitive tasks can be represented as follows:

$$t_{-i1} = \sqrt{p_{-A1}} * A_{-i} + \sqrt{p_{-c1}} * C_{-i} + \sqrt{p_{-e1}} * e_{-i1} + \sqrt{p_{-o1}} * o_{-i1} \tag{20}$$

$$t_{-i2} = \sqrt{p_{-B2}} * B_{-i} + \sqrt{p_{-c1}} * C_{-i} + \sqrt{p_{-e2}} * e_{-i2} + \sqrt{p_{-o1}} * o_{-i1} \tag{21}$$

Both of these cognitive tasks measure C. They also share o which includes sources of reliable variance not included in T. It should also be noted that each basic cognitive task measures an ability not measured by the other. t_{-i1} includes ability A and t_{-i2} includes ability B. All measures are assumed to be in z-score form.

As for Models I and III, the development of Model II begins with the deviation score definition of the correlation between a standardized test and a basic cognitive task:

$$r_{-T11} = \sum(x_{-T} * x_{-i}) / (N S_{-T} S_{-i}) \quad (22)$$

Substituting the raw score equivalents because all scores are in z-score form gives:

$$= \sum(T_{-i} * t_{-i1}) / (N S_{-T} S_{-i}) \quad (23)$$

Because $S_{-T} = S_{-i} = 1.00$ the expression can be written:

$$= \sum(T_{-i} * t_{-i1}) / N \quad (24)$$

Expanding the raw scores with the definition previously given yields:

$$= \sum(\sqrt{P_{-A}} * A_{-i} \dots \sqrt{P_{-E}} * E_{-i})(\sqrt{p_{-A1}} * A_{-i} \dots \sqrt{p_{-e1}} * e_{-i1}) / N \quad (25)$$

Multiplying the two expressions, nearly all terms drop out except:

$$= \sum(\sqrt{P_{-A}} * A_{-i} * \sqrt{p_{-A1}} * A_{-i} + \sqrt{P_{-C}} * C_{-i} * \sqrt{p_{-C1}} * C_{-i}) / N \quad (26)$$

which can be written as:

$$= \sum(\sqrt{P_{-A}} * \sqrt{p_{-A1}} * A_{-i2} + \sqrt{P_{-C}} * \sqrt{p_{-C1}} * C_{-i2}) / N \quad (27)$$

and because all weights are constants, the expression can be simplified to:

$$= \sqrt{P_{-A} p_{-A1}} * (RA_{-i2}) / N + \sqrt{P_{-C} p_{-C1}} * (RC_{-i2}) / N \quad (28)$$

Because the sum of z-scores squared is equal to N , the expression becomes:

$$= \sqrt{P_{-A} p_{-A1}} + \sqrt{P_{-C} p_{-C1}} \quad (29)$$

Thus, the correlation between a basic cognitive task and a standardized test according to Model III is the sum of the square root of the product of the weights of each ability measured by the basic cognitive task.

By similar development, it can be shown that

$$r_{-t1t2} = \sqrt{p_{-C1} p_{-C2}} + \sqrt{p_{-O1} p_{-O2}} \quad (30)$$

The correlation between two basic cognitive tasks is a function of the weights of the abilities both tasks measure in common and any other common variance they share (p_{-O1} and p_{-O2}) but that is not measured by g .

What is the lower limit of the correlation between cognitive tasks given that the two tasks involved have a given correlation with g ? Because Model II defines abilities as being independent, it is possible that two tasks could measure different sets of abilities and be uncorrelated. So a lower limit on the correlation between two sets of basic cognitive tasks according to Model II will always be zero.

SUMMARY AND COMPARISON OF EVIDENCE FOR MODELS I AND III VS. MODEL II

Under the assumptions of Model I and III, it can be shown that the lower limit for the correlation between cognitive tasks (r_{T1T2}) is the product of the correlation of each cognitive task with the standardized test representing g ($r_{T1} * r_{T2}$). For Model II, the lower limit of the correlations between cognitive tasks is 0 and this lower limit is not dependent on the correlation of the cognitive tasks with a standardized test. Figure 5 shows the lower limits of the correlations between cognitive tasks based on a given correlation between the cognitive task and the standardized test. For example, if two cognitive tasks each correlate with a standardized test .40, the lower limit of the correlation between them according to Model I and III would be .16 and under Model II it would be 0.

The different intertask correlations predicted by Models I and III and Model II for a given level of correlation between a cognitive task and standardized test can be used to discriminate among the theoretical positions. At the extreme, if two tasks can be shown to correlate with g or a standardized test representing g but not be correlated with each other, Model II must be correct. On the other hand, if the distribution of correlations of cognitive tasks is predicted by the product of their individual correlations with g , then Models I and III are supported. This simple difference between theoretical positions has a number of implications, but before discussing those, another issue needs to be considered.

Explanations of Low Relationships

Earlier it was mentioned that typically correlations between standardized tests and basic cognitive tasks have been found to be low. This is entirely consistent with Model II. Consider the formula for a correlation between overlapping elements:

$$r = 1/\sqrt{(1+m)(1+k)} \quad (31)$$

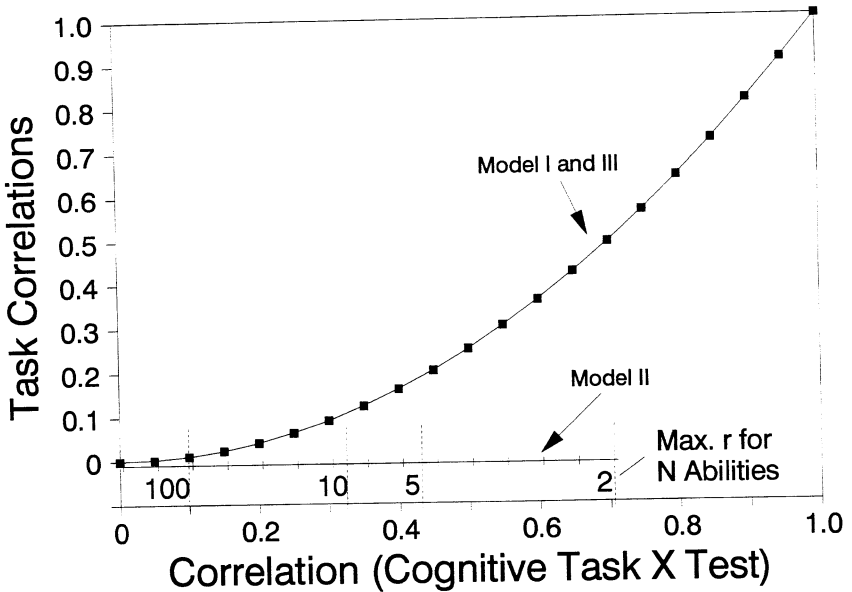


Figure 5. The correlation among cognitive tasks as a function of the correlation of the tasks with a standardized test measuring general intelligence. The maximum correlation for n abilities for Model II is also shown. For example, if there were five abilities composing g , each would correlate .45 ($\sqrt{.20}$) with the standardized test of general intelligence.

where l is the number of common elements, m is the number of elements unique to Set 1, and k is the number of elements unique to Set 2. Let Set 1 represent a standardized test which measures 10 independent abilities. Let Set 2 represent a cognitive task which measures only one of the 10 independent abilities and nothing else. Then l , the number of common elements, is 1; m , the number of elements unique to set 1 is 9; and k is 0. Substituting these values into Equation 31:

$$r_{-T1} = 1/\sqrt{(1+9)(1+0)} = 1/\sqrt{10} = \sqrt{10} = .316 \quad (32)$$

So if intelligence were composed of 10 independent abilities and cognitive tasks each measured only one ability, the average expected correlation for perfectly reliable tests would be .32. The expected correlation is dependent on the number of independent abilities:

$$r_{-T1} = 1/\sqrt{NA} \quad (33)$$

where NA is the number of independent abilities. Therefore, the number of abilities determines the expected correlation between standardized tests and basic cognitive tasks if basic cognitive tasks are measuring only a single ability.

Figure 5 shows the expected maximum correlation between basic cognitive tasks and standardized tests for given numbers of abilities. It should be noted that the larger the number of abilities, the more difficult Model II will be to discriminate from Models I and III. Somewhere above 100 basic abilities, the models may be theoretically indiscriminable. They are practically indistinguishable if there are more than 100 abilities on the basis of the evidence being discussed here. But for smaller numbers of abilities, they can be discriminated.

There appears to be no simple explanation of the low correlations between basic cognitive tasks and standardized tests that can be generated from the assumptions of Models I and III. In fact, the standard explanation would be that the cognitive tasks were not *g*-loaded. Since low correlations would indicate the absence of *g*, there would be little theoretical interest in these tasks by proponents of Models I and III. They would want to find tasks that had high correlations (were heavily *g*-loaded). Even more importantly, tasks that had low correlations would be regarded as having little to do with intellectual functioning. So most basic cognitive tasks measuring such things as memory, learning, and perception would be viewed as having little or nothing to do with intelligence since such tasks are typically found to have low correlations.

Zero Correlation

As was discussed earlier, finding that two tasks are not correlated yet correlate with a measure of *g* is important support for Model II and would require that Models I and III be rejected. If basic cognitive tasks were found that had a zero correlation with each other but still correlated with standardized measures of intelligence (or *g*), there would be other, secondary results that would also be found that are discussed below. These secondary results might be easier to test and have important implications of their own.

But before considering secondary results, it is reasonable to ask if there is any evidence that would suggest that correlations between cognitive tasks that correlate with measures of intelligence could be zero. The prevailing belief is in *positive manifold*; that is, that all mental tests display positive correlations. Guilford (1964) analyzed 13 correlation matrices which had been obtained from batteries of intellectual tests. These batteries contained over 7,000 correlations. It is interesting to note that the median correlation in these matrices fell between .2 and .3, somewhat lower than would be expected even from subtests of most intellectual tests. From the low median correlation, it appears that the tests that Guilford analyzed were more like basic cognitive tasks than standardized tests of intellectual functioning. In any event, somewhere between 17% and 24% of the correlations could be considered zero values. Guilford

concludes that for purposes of factor analyses this is a sufficient number of zero correlations to establish orthogonal simple structure. Regardless of their implications for factor analysis, Guilford's results suggest that it is possible to find mental tests that have zero correlations with each other but still measure intelligence.

The finding of zero correlations among mental tests is theoretically a very important finding. Models I and III could not account for this finding. The correlation of tests with each other is directly related to their g loading. On the other hand, Model II predicts that such zero correlations should be obtainable. A pure measure of an underlying process should have a zero correlation with a pure measure of any other process. Guilford's finding of zero correlations among mental tests is a critical piece of information. A critical test of the validity of Model II is the ability to find basic cognitive tasks that predict general intelligence but are uncorrelated with each other. Failure to find such correlations is not support for Models I and III. That is because Model II could explain the lack of such correlations as a failure to measure pure processes.

Size of First Principal Component

An important piece of support for Models I and III has been the percentage of variance accounted for by the first principal component. For batteries of standardized tests of intelligence, the percentage of variance accounted for is generally between 40% and 80%. This is an important finding because the magnitude of the first principal component is a rough representation of the average correlation among the tests in the battery. The lower the correlation among the tests, the smaller the size the first principal component would be. Model II would predict that, if it were possible to obtain good representations of the basic processes underlying ability, the size of the first principal component would be substantially reduced. Models I and III are not clear about what such a finding would mean. Because low correlations among tests probably indicate that the tests are not very reliable measures of g , a small first principal component means that the tests in the matrix are simply bad measures and poor predictors of g .

Detterman et al. (1992; see also Detterman, 1992) have demonstrated that the amount of variance accounted for by the first principal component is substantially reduced for basic cognitive tasks from levels that would be expected from standardized intelligence tests. This is not really surprising because the amount of variance accounted for by the first factor depends on the average correlation among the variables in the matrix. These correlations are lower among basic cognitive tasks than among standardized intelligence tests. Interestingly, lower IQ persons show higher correlations among basic cognitive tasks than higher IQ persons (Detterman & Daniel, 1989). This finding would

certainly not be anticipated by Models I and III, though it can be explained by Model II.

Multiple Regression

If a set of measures from basic cognitive tasks were used to attempt to predict factor scores representing g or a standardized test of intelligence, what sorts of results would be expected according to each of the theoretical positions? According to Models I and III, measures from basic cognitive tasks represent g . If the correlations are low it is only because the tasks do not provide very good measures of g . Combining such measures in a multiple regression is exactly like increasing a test's length. No new sources of variance are added by adding additional measures to the multiple regression; the measure of g is being made more reliable. Therefore, according to Models I and III, when measures are combined in a multiple regression to predict g , the increase in variance accounted for would parallel the increase that would be found by increasing test length.

In marked contrast, the prediction which can clearly be made from Model II is that, to the extent the measures are independent measures of the separate processes making up g , when entered into a multiple regression the measures should be additive. That is, because each measure would represent a separate, nonoverlapping source of variance, these sources should simply add together. If the measures obtained from basic cognitive tasks were not purely representative of the underlying independent cognitive abilities, the measures would be correlated with each other. This, of course, would mean that each measure would be a poorer representation of a single underlying ability and so, each measure would contribute less to the regression equation.

The two theoretical positions make markedly different predictions with respect to multiple regression. Models I and III predict that as the correlation among basic cognitive tasks gets higher, the multiple correlation of those tasks with a traditional measure of g will increase. On the other hand, Model II predicts that the highest multiple correlation would be obtained when the correlation between cognitive tasks is zero so long as the basic tasks correlate with g . In other words, Models I and III predict that the size of the multiple correlation will be a direct function of the percentage of variance accounted for by the first principal component, but Model II predicts the two measures will vary inversely (particularly as the average correlation among basic cognitive tasks approaches zero).

Although there is no definitive data bearing on this issue, Detterman et al. (1992) did find that basic cognitive tasks could be used to predict IQ or g about as well as IQ tests can predict each other. This result occurred even when the correlation between measures entered into the equation was minimized.

Basic Cognitive Tasks as Predictors of g

Kranzler and Jensen (1991a, 1991b, 1993; see also comments by Carroll, 1991a, 1991b, 1993) have used a clever methodology to test if Model II is a better explanation of g than Models I and III. First, they derived a measure of g in the usual way using standard tests of intelligence. Second, they factor analyzed a set of basic cognitive tasks. They argued that if g is a unitary construct as suggested by Models I and III then the g from basic cognitive tasks should be the same as the g from standard intelligence tests. In other words, g from the standard tests (the first principal factor) should be correlated only with the first principal factor from the basic cognitive tasks. Other factors from the basic cognitive tasks should not correlate with g . More than one factor from the basic cognitive tasks correlating with g would provide support for Model II.

Kranzler and Jensen found that three additional factors beyond the first were significantly correlated with g . Even after making some radical alterations suggested by Carroll, one additional factor beyond the first was significantly correlated with g . Kranzler and Jensen have provided convincing evidence supporting Model II.

Biological Systems

Thompson, Crinella, and Yu (1990) have provided biological evidence in support of Model III. They replicated the kinds of ablation studies done by Lashley (1929). The advances in anatomical knowledge allowed more precise structural placement of lesions. Subjects were also given a larger battery of tests than in the original Lashley experiments. In all, 424 rats were tested, a sample large enough for the sophisticated analyses that were conducted.

The results from this huge study were analyzed with the hope of finding a general learning system that might be analogous to g in humans. Instead, the authors conclude, "One fact becomes abundantly clear: *Almost every brain structure implicated in psychometric g participates in a different neural mechanism* (Thompson et al., p. 154; emphasis in original). Their results suggest that psychometric g is composed of a small number of basic neural mechanisms. This finding is obviously strong support for Model II.

Idiot Savants and Other Exceptional Abilities

One phenomena that Models I and III have difficulty explaining is the existence of exceptional abilities such as those obtained by idiot savants, exceptional memorizers, lightning calculators, calendar calculators, hyperlexia, and other unusual abilities. Such abilities can only be accounted for by practice or by assuming that the abilities are not associated with general intelligence.

Conners (1992) recently reviewed literature on outstanding abilities and concluded that such explanations were not convincing. While data on persons with these unusual abilities is scanty, results seem to be most consistent with Model II.

Critical Evidence for Models I and III

The sources of evidence that have been discussed so far would provide critical support for Model II but not necessarily for Models I and III. This is because findings in support of Model II are inconsistent with Models I and III but evidence in support of Models I and III might be explained, with further analysis, by Model II. Results supporting Models I and III are not necessarily inconsistent with Model II. Is there critical evidence that would support Models I and III?

There is probably no critical evidence to support Models I and III, but there is a type of evidence that would provide strong tentative support for the position. If a measure from basic cognitive tasks was found that correlated highly with standard measures of g and could not be explained in terms of more basic cognitive measures, this measure would provide strong support for Models I and III. Coupled with parsimony, the fact that a single variable provides a more appealing explanation of a phenomenon than more complex explanations of the same phenomenon, this evidence would provide strong support for Models I and III.

In the absence of evidence supporting either of the theoretical positions, it could well be argued that on the basis of parsimony alone, Models I and III should be the explanatory choice. A consideration of some of the phenomena to be explained by these theories is in order to determine which seems the most appropriate choice.

SUMMARY AND CONCLUSION

Three models of intelligence have been developed which cover the theoretical possibilities of accounting for g . Aspects of these models have been formally developed. It was shown that Models I and III were functionally equivalent. Both regard g as essentially a single thing. Model II postulates that g is composed of a finite number of independent abilities.

Evidence that could bear on the discrimination of Model II from Models I and III was summarized. The bulk of evidence presented supported Model II. More importantly, the kinds of evidence needed to support alternative theoretical positions were detailed. There are a number of sources of critical evidence for the different positions. The general conclusion is, then, that it will be pos-

sible to ultimately determine if g is better accounted for by Models I and III or by Model II.

Elsewhere (Detterman, 1982, 1984, 1986, 1987) I suggested a systems theory of intelligence that falls in the Model II class of models. The system model that I proposed is a more restricted theory than the general Model II. I postulated that some mental processes will be more important than others and that some will be more centrally important to system functioning. All of the evidence presented in support of the Model II class of models is also consistent with the system theory that I proposed. This agreement is a logical necessity since the theory I proposed is a subset of all possible Model II models and the evidence presented was in support of the general class of Model II models. However, it would be possible to disconfirm the systems theory that I proposed without affecting the general status of Model II models.

REFERENCES

- Binet, A., & Henri, V. (1896). La psychologie individuelle. *Individual Psychology L'Annee Psychologique*, 2, 411–465.
- Binet, A., & Simon, T. (1908). New methods for the diagnosis of intellectual level of the abnormal. *Methodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. L'Annee Psychologique*, 11, 191–244.
- Carroll, J. B. (1980). *Individual difference relations in psychometric and experimental cognitive tasks*. Chapel Hill, NC: L. L. Thurstone Psychometric Laboratory, University of North Carolina.
- Carroll, J. B. (1991a). No demonstration that g is not unitary, but there's more to the story: Comment on Kranzler and Jensen. *Intelligence*, 15, 423–436.
- Carroll, J. B. (1991b). Still no demonstration that g is not unitary: Further comment on Kranzler and Jensen. *Intelligence*, 15, 449–453.
- Carroll, J. B. (1993). The unitary g problem once more: On Kranzler and Jensen. *Intelligence*, 17, 15–16.
- Cattell, J. M. (1890). Mental tests and measurements. *Mind*, 15, 373–381.
- Conners, F. A. (1992). Special abilities of idiots savants, hyperlexic children, and phenomenal memorizers: Implications for intelligence theory. In D. K. Detterman (Ed.), *Current topics in human intelligence, Vol. 2: Is mind modular or unitary?* (pp. 187–277). Norwood, NJ: Ablex.
- Detterman, D. K. (1979). Memory in the mentally retarded. In N. R. Ellis (Ed.), *Handbook of mental deficiency* (2nd ed., pp. 727–760). Hillsdale, NJ: Erlbaum.
- Detterman, D. K. (1982). Does 'g' exist? *Intelligence*, 6, 99–103.
- Detterman, D. K. (1984). Understand cognitive components before postulating meta-components, etc.: Part II. *Behavioral and Brain Science*, 7, 289–290.
- Detterman, D. K. (1986). Human intelligence is a complex system of separate processes. In R. J. Sternberg & D. K. Detterman (Eds.), *What is intelligence? Contemporary viewpoints on its nature and definition*. (pp. 57–61) Norwood, NJ: Ablex.

- Detterman, D. K. (1987). Theoretical notions of intelligence and mental retardation. *American Journal of Mental Deficiency, 92*, 2–11.
- Detterman, D. K. (1992). Mopping up: The relation between cognitive abilities and intelligence. *American Journal of Mental Retardation, 97*, 295–301.
- Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. *Intelligence, 13*, 349–359.
- Detterman, D. K., Mayer, J. D., Caruso, D. R., Legree, P. J., Conners, F., & Taylor, R. (1992). Assessment of basic cognitive abilities in relation to cognitive deficits. *American Journal of Mental Retardation, 97*, 251–286.
- Ellis, N. R. (1963). The stimulus trace and behavioral inadequacy. In N. R. Ellis (Ed.), *The handbook of mental deficiency* (pp. 134–158). New York: McGraw-Hill.
- Ellis, N. R. (1970). Memory processes in retardates and normals. In N. R. Ellis (Ed.), *International review of research in mental retardation* (Vol. 4, pp. 1–32). New York: Academic Press.
- Eysenck, H. J. (Ed.). (1982). *A model for intelligence*. New York: Springer-Verlag.
- Galton, F. (1883). *Inquiries into human faculty and its development*. London: Macmillan.
- Garnett, J. C. M. (1919a). *Proceedings of the Royal Society, Section A, 96*, 91–100.
- Garnett, J. C. M. (1919b). General ability, cleverness and purpose. *British Journal of Psychology, 9*, 345–366.
- Garnett, J. C. M. (1920). The single general factor in dissimilar mental measurements. *British Journal of Psychology, 10*, 241–258.
- Guilford, J. P. (1964). Zero correlations among tests of intellectual abilities. *Psychological Bulletin, 61*, 401–404.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Hart, B., & Spearman, C. (1912). General ability, its existence and nature. *British Journal of Psychology, 5*, 51–84.
- Humphreys, L. G. (1979). The construct of general intelligence. *Intelligence, 3*, 105–120.
- Hunt, E., Lunneborg, C., & Lewis, J. (1974). What does it mean to be high verbal? *Cognitive Psychology, 7*, 194–227.
- Jensen, A. R. (1979). *g*: Outmoded theory or unconquered frontier? *Creative Science and Technology, 2*, 16–29.
- Jensen, A. R. (1980). Chronometric analysis of mental ability. *Journal of Social and Biological Structures, 3*, 103–122.
- Jensen, A. R. (1982a). Reaction time and psychometric *g*. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 93–132). New York: Springer-Verlag.
- Jensen, A. R. (1982b). The chronometry of intelligence. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1). Hillsdale, NJ: L Erlbaum.
- Jensen, A. R. (1984). Test validity: *g* versus the specificity doctrine. *Journal of Social and Biological Structures, 7*, 93–118.
- Jensen, A. R., & Munro, E. (1979). Reaction time, movement time, and intelligence. *Intelligence, 3*, 121–126.
- Jensen, A. R., & Weng, L. -J. (in press). What is a good *g*? *Intelligence*.

- 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.
- Johnson, R. C., McClearn, G. E., Yuen, S., Nagoshi, C. T., Ahern, F. M., & Cole, R. E. (1985). Galton's data a century later. *American Psychologist*, *40*, 875–892.
- Kranzler, J. H., & Jensen, A. R. (1991a). The nature of psychometric *g*. *Intelligence*, *15*, 397–422.
- Kranzler, J. H., & Jensen, A. R. (1991b). Unitary *g*: Unquestioned postulate or empirical fact. *Intelligence*, *15*, 437–448.
- Kranzler, J. H., & Jensen, A. R. (1993). Psychometric *g* is still not unitary after eliminating supposed “impurities”: Further Comment on Carroll. *Intelligence*, *17*, 11–14.
- Lashley, K. S. (1929). *Brain mechanisms and intelligence*. Chicago: University of Chicago Press.
- Mackie, J. (1928). The sampling theory as a variant of the two factor theory. *Journal of Educational Psychology*, *19*, 614–621.
- Maxwell, A. E. (1972). Factor analysis: Thomson's sampling theory recalled. *British Journal of Mathematical and Statistical Psychology*, *25*, 1–21.
- Sharp, S. E. (1898). Individual psychology: A study in psychological method. *American Journal of Psychology*, *10*, 329–391.
- Spearman, C. E. (1904). “General intelligence” objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.
- Spearman, C. E. (1927). *The abilities of man: Their nature and measurement*. London: Macmillan.
- Sternberg, R. J., & Salter, W. (1982). Conceptions of intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 3–28). New York: Cambridge University Press.
- Sternberg, R. J., & Spear, L. C. (1985). A triarchic theory of mental retardation. In N. R. Ellis & N. Bray (Eds.), *International review of research in mental retardation* (Vol. 13, pp. 301–326). New York: Academic Press.
- Thompson, R., Crinella, F. M., & Yu, J. (1990). *Brain mechanisms in problem solving and intelligence*. New York: Plenum.
- Thomson, G. H. (1916). A hierarchy without a general factor. *British Journal of Psychology*, *18*, 68–76.
- Thomson, G. H. (1919). On the cause of hierarchical order among correlation coefficients. *Proceedings of the Royal Society, A*, *95*, 400–408.
- Thomson, G. H. (1927). A worked-out example of the possible lineage of four correlated variables on the sampling theory. *British Journal of Psychology*, *18*, 68–76.
- Thomson, G. H. (1935). On complete families of correlation coefficients and their tendency to zero tetrad-differences: Including a statement of the sampling theory of abilities. *British Journal of Psychology*, *26*, 63–92.
- Thomson, G. H. (1939). *The factorial analysis of human ability*. London: University of London Press.
- Thorndike, E. L. (1927). *The measurement of intelligence*. New York: Teachers College, Columbia University.
- Thurstone, L. L. (1935). *Vectors of mind*. Chicago: University of Chicago Press.

- Thurstone, L. L. (1938). *Primary mental abilities. Psychometric Monographs* (No. 1). Chicago: University of Chicago Press.
- Tuddenham, R. D. (1963). The nature and measurement of intelligence. In L. Postman (Ed.), *Psychology in the making* (pp. 469–523). New York: Knopf.
- Undheim, J. O., & Horn, J. L. (1977). Critical evaluation of Guilford's Structure-of-Intellect theory. *Intelligence, 1*, 65–81.
- Vernon, P. A. (Ed.). (1987). *Speed of information-processing and intelligence*. Norwood, NJ: Ablex.
- Whipple, G. M. (1910). *Manual of mental and physical tests*. Baltimore: Warwick & York.
- Wissler, C. (1901). The correlation of mental and physical tests. *Psychological Review Monograph Supplement, 3*(6).
- Zeaman, D., & House, B. J. (1963). The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), *Handbook of mental deficiency* (pp. 159–223). New York: McGraw-Hill.

Chapter 6

A Biological Theory of Intelligence

H. J. Eysenck

*Institute of Psychiatry
University of London*

THE DEFINITION OF INTELLIGENCE

As I pointed out almost 40 years ago, intelligence testing is a technology without a scientific theory to support it (Eysenck, 1953). If we believe with K. Levin that there is nothing as practical as a good theory, we ought to be sorry for this lack. Equally, science generally proceeds best when working within a paradigm based on a general theoretical conception of the particular field in question (Popper, 1935); this enables fruitful “normal science” to go on and eliminate the many anomalies which usually characterize even advanced sciences (Kuhn, 1970, 1974). What is disappointing is not only the absence of a natural-science theory of intelligence, but the lack of concern about its absence, and the failure of psychologists to even try to work toward such a theory.

There are many problems in defining just what a proper theory in science is, and what it does (Suppe, 1974), but broadly speaking such a theory would obviously have to (a) explain and pull together the major facts known in the field, and (b) make testable predictions about hitherto unknown facts. There are many interesting philosophical discussions of the differences between explana-

tion and understanding, with the oft-repeated assertion that explanations of human actions cannot be reduced to simple causality (Wright, 1971). Many others doubt the possibility of producing a genuine theory of intelligence, to take its place beside the theories of the "hard" sciences. There is little point in philosophical debates; even in science, that which was declared impossible only the day before has often been accomplished the day after. J Muller, in his famous *Handbuch der Physiologie*, predicted that we would probably never know the velocity of the nervous impulse, only a few years before his erstwhile pupil Helmholtz did measure it (Boring, 1957). I will not deal with philosophical contentions that such a theory must be impossible to produce, but I will cite detailed comparisons between attempts to produce theories in physics, such as the theory of gravitation, and the production of a theory of intelligence, comparing, as it were, G with g . Such comparisons sharpen our understanding of what theories can and cannot do, and how they originate and develop.

Philosophical doubts bite deep, and it is often asserted that searching for a theory of intelligence is useless because "intelligence" does not exist. Of course, "intelligence" does not exist, in any meaningful physical sense; it is a scientific concept, like gravitation, or heat, or society. Concepts do not exist in the sense that stones, or pigs, or stars "exist." They are human inventions to bring order into the blooming, buzzing confusion of everyday life events, and they are more or less useful or useless (Eysenck, 1988). Existence is irrelevant; concepts explicitly or implicitly embody theories which can be verified or disconfirmed ("falsified" in Popper's language). That is their main function in science.

It may be objected that there is no agreed concept of intelligence, but that there is a great variety of concepts mutually contradicting each other, as shown by Sternberg and Detterman (1986), a result not dissimilar to that of an earlier effort at definition published in the *Journal of Educational Psychology* (1921). Intelligence is there defined in terms of adaptive behavior, learning, problem solving, knowledge acquisition, goal achievement, planning, information coding and processing, attention, symbol manipulation, judgment, attribution, a hodgepodge of cognitive capacities, a repertoire of intellectual knowledge and skills, understanding, purposive thinking, apprehension of experience, mental self-government, and many more. Such variety has often been used to castigate users of the term; if there is such disagreement, psychologists clearly do not know what the term means. Snyderman and Rothman (1987, 1988) found a good deal of agreement among the 600-plus experts consulted; abstract thinking or reasoning was checked by 99.3% of respondents as important, problem-solving ability by 97.7%, and capacity to acquire knowledge by 96.0%. Does this help?

I have suggested that the whole approach is misconceived (Eysenck, 1988). All the definitions of the concept deal in fact with *effects* the concept may have; intelligence is that which *enables* us to think abstractly, to solve prob-

lems, to acquire knowledge, to adapt our behavior to circumstances, to achieve goals, to plan, to code information, to manipulate symbols, to perform judgment, to understand, to apprehend experience, and so forth. Whatever it is, it is a *capacity* to carry out cognitive tasks successfully; it is a *dispositional* variable enabling the possessor to do all the things mentioned more or less successfully.

In the same way, physicists do not define gravitation in terms of the apple falling on Newton's head, the shape of the planets, the occurrence of tides or black holes, the equatorial bulge of the earth, planetary movements, the laws of gunnery, or the formation of galaxies. Gravitation is what makes all these things happen, but it cannot be *identified* with any of them. What we are looking for is something underlying the variety of applications, both in regard to G or g —gravitation in physics, intelligence in psychology.

The search for this underlying reality has two phases, and it is important to remember this. The first is *descriptive* and taxonomic, the second is *causal*. Consider gravitation. Newton succeeded in writing a formula which provided a general description unifying a large number of phenomena, involving the product of the masses of any two bodies interacting, and the square of the distance between them. The description turned out to be surprisingly successful, but it did not constitute a theory; all Newton managed was the theory of action at a distance, which he himself recognized as absurd, and which certainly did not *explain* the success of his laws. In fact there is *still* no accepted theory of gravitation, 300 years after Newton. There are several attempts at a theory, but no agreed theory. We have Einstein's field theory explaining G as a distortion of space-time geometry. We have the quantum mechanics theory of particle interaction (gravitons). And we finally have Sakharov's theory of gravity as a long-range Casimir force, that is, the postulation that gravity might be an effect brought about by changes in the zero-point energy of the vacuum, due to the presence of matter. No agreement is in sight on a reconciliation of these quite different conceptions. Where physicists have failed after 300 years to reach agreement on a general theory of gravitation, should we be surprised that psychologists have equally failed in their possibly more difficult endeavor, after at most 100 years?

Of course intelligence and gravitation are concepts which humankind has been familiar with for thousands of years, in a very practical sense. We know how gravitation works; how else could we throw spears, kick a football accurately, run and jump, build houses, and dig wells? The attractive force of the earth is as familiar to us, almost from birth, as is the variable ability of others to understand, to learn, and to solve problems. Equally, from the earliest days there has been a bifurcation of meanings and uses of the term *intelligentia*, defined in Latin dictionaries (Lewis & Short, 1945) either as *understanding* (capacity) or *knowledge* (achievement). Similarly, the *Short Oxford English Dictionary* (Little, Fowler, & Coulson, 1947) defines *intelligence* as either "faculty of understanding; understanding as a quality admitting of degree;

achievement of mental apprehension" (capacity), or as "information" (knowledge). The verb *intelligere* similarly may mean "to understand, comprehend," or "to have an accurate knowledge of; to distinguish" (Lewis & Short, 1945). The *Thesaurus Linguae Latinae* gives countless examples of the use of these terms by Cicero and other writers, but they seldom go beyond these two major applications.

This distinction between *capacity* and *achievement* is reflected in modern theories of fluid (g_f) and crystallized (g_c) ability (Cattell, 1963; Horn & Cattell, 1966, 1982). The former is conceptualized as abstract, essentially nonverbal, relatively culture-free mental efficiency, while the latter is viewed as consisting primarily of acquired skills and knowledge, and thus strongly dependent on cultural exposure. This distinction is well supported, and there is evidence of a causal nature that g_f determines later g_c , at least in middle-class children (Schmidt & Crano, 1974). It is tempting to think of g_f as more strongly determined by genetic factors, and g_c by experiential ones, but Cattell's (1971) own results show little difference, heritability being 0.77 for g_f and 0.73 for g_c . Similarly, Shields (1962), in a study of MZ twins, found little difference in the intraclass correlation for dominoes (a g_f test) and vocabulary (a g_c test) for either twins brought up together (0.71 and 0.74) or in separation (0.70 and 0.74). Possibly differences would have been more apparent in a less homogeneous culture. Cattell's (1980) latest study gives 45% for the heritability of g_c and 58% for g_f , using his general population values; this is more in line with expectations. Horn (1985) failed to find g_f determined more strongly by genetic factors than g_c ; this issue is not settled.

However that may be, the distinction in meaning and definition is important, with capacity clearly constituting the more fundamental variable, even though acquired knowledge may be socially more important, and certainly easier to measure. There is some correspondence between g_f and potential energy, and g_c and kinetic energy, although the analog should not be taken too far. Any theory of intelligence would certainly have to find a more central place for g_f rather than g_c , and as we shall see, this is indeed the case.

It will also be obvious that while the acquisition of knowledge depends in part on g_f , it also depends on other factors, such as motivation, personality, opportunity, and so on, and can hence never be a *fundamental* part of psychological science, because of its compound nature. Alexander (1935) showed the importance of persistence in school achievement, and there is ample evidence for the involvement of personality (Eysenck, 1973a, 1973b, 1978); opportunity, luck, and other factors like socioeconomic status no doubt also contribute. Hence, there can be no scientific theory of g_c as such, but only theories of the diverse factors contributing to it. Under certain circumstances, such as equal opportunities of schooling, advancement, and working, g_c may nevertheless give us a good approximation to a measure of g_f . The usual correlation between g_f and g_c is around 0.60, but it can be higher in more homogeneous groups (and lower in less homogeneous groups); indeed the correlation

between g_f and g_c might be regarded as a measure of cultural homogeneity. (Homogeneity with respect to *ability* would of course have the opposite effect, through restriction of range.)

A proper definition of intelligence would not have to stray too far from its accepted meaning in popular terminology (Sternberg, Conway, Ketron, & Bernstein, 1981). It relates to *cognitive* behavior (learning, problem solving, abstract thinking, and reasoning), appears as a *variable* ability differentiating people on a graded scale, and must be regarded as a *capacity* for successful cognition. If we add a strong genetic predisposition (Eysenck, 1979; Plomin, De Fries, & McClearn, 1990; Vernon, 1979) and the universal extension of the capacity notions over all cognitive tasks (Gustafsson, 1984; Jäger & Hoemann, 1983; Undheim & Gustafsson, 1987), we have the following definition: Intelligence is the general, graded cognitive capacity of a person, largely but not exclusively determined by genetic factors. Problems arising from this definition will be discussed in the next section.

THEORY AT THE TAXONOMIC OR DESCRIPTIVE LEVEL

Scientific theories have two not unconnected territories to cover. One, the more fundamental and primitive, is the taxonomic or descriptive; the other is the causal (Eysenck, 1969; Sokal & Sneath, 1963). Before we can create causal theories, we must have some idea of what it is that we wish to account for. Taxonomy must precede theories to explain the findings of taxonomy; we must classify animals before we can arrive at a theory of evolution. The point is universally recognized in science. Even in astronomy, where we seldom think of taxonomy, it forms an indispensable basis for more experimental and empirical work (Jaschek & Jaschek, 1990). Thus it is basic to any evolutionary account of stellar development, as shown for instance in the famous Hertzsprung-Russell diagram (Hoyle, 1962).

Similarly, we must classify the capacity or capacities we have postulated. Do we have to find one or several causes for one universal g , or for a multitude of capacities (Detterman, 1982; Gardner, 1984; Guilford, 1967, 1981a, 1981b; Guilford & Hoepfner, 1971; Maxwell, 1972; Neisser, 1979; Thomson, 1939; Thorndike, Bregman, Cobb, & Woodyard, 1926)? Non- g theories have certainly attracted much attention, as have attempts to account for the facts by attributing the "positive manifold" and the low rank of correlated matrices to "bonds" of one kind or another, but these attempts have not been successful. Guilford's theory has been criticized by Undheim and Horn (1977), Eysenck (1979), Horn and Knapp (1973), and many others; the model cannot be sustained in view of the many damaging faults that have been found.

Evaluation of the various theories postulating some variety of "bonds," not usually very carefully defined, must be equally negative. Loevinger (1951)

argued that Thomson's theory makes "no assertions to which evidence is relevant" (p. 595), so that the theory would be untestable. Eysenck (1987a) has argued that testable deductions are possible and has provided evidence against the theory. It is no longer tenable, and certainly does not account for the facts it is supposed to explain.

On the whole, the case originally made by Spearman (1927) for g as an explanation of the "positive manifold" and the low rank of empirically observed matrices is difficult to argue against. Gustafsson (1984), Jäger (1967, 1984), Jensen (1984, 1986b), Marshalek, Lohman, and Snow (1983), Nagoshi and Johnson (1986), Snow (1980), Snow, Lohman, Marshalek, Yalow, and Webb (1977), and many others have shown by the use of factor analysis, confirmatory factor analysis, and multidimensional scaling that descriptively a hierarchical model along the lines pioneered by Burt (1909, 1940), Vernon (1950), and Eysenck (1939), with g at the top, corresponds best with reality and succeeds in bringing together Spearman's (Spearman & Jones, 1950) and Thurstone's (Thurstone & Thurstone, 1941) at first antagonistic conceptions.

The argument goes well beyond this explanation of positive manifold and low ranks in terms of g ; it leads on to a prediction of which tests would have highest loadings on the general factor, or alternatively, which tests would occupy the central spot in a multidimensional scaling analysis. According to Spearman's theory of noegenesis, such tests would involve the discovery of relations and correlates, that is, involve abstract problem solving, analysis, and the determination of rules. In fact, such tests do show the highest loadings on g , and occupy a central space in smallest space analysis. Crucial perhaps is the fact that Raven's Matrices occupy a central position in the two-dimensional multidimensional scaling model (Marshalek et al., 1983). A "bonds" theory would predict that a high-loading g test would require "all the components of the mind" (Maxwell, 1972, p. 4). The total WAIS score is very close to the Raven score in the analysis, and might be thought to fulfill this condition, but the Raven could not by any stretch of the imagination be thought to do so. This type of analysis in terms of Guttman's (1954) *radex* would seem to give results incompatible with any form of the "bonds" theory, and support Spearman's (1923) interpretation. It is important to realize that our hypotheses concerning the (descriptive) nature of intelligence can be supported or refuted by the statistical analysis of matrices of intercorrelations between tests; the Raven was explicitly constructed following Spearman's laws of noegenesis.

The evidence now seems very strong that not only must we postulate some kind of hierarchical model of intelligence, but that much the greatest part of the variance belongs to g , and that $g = g_f$ (Gustafsson, 1984; Snow et al., 1977; Snow, Kyllonen, & Marshalek, 1984). Confirmatory factor analysis and multidimensional scaling have now clarified the picture sufficiently to allow us to try and go beyond the descriptive phase. There are, of course, many objections

to the use of factor analysis (e.g., Revelsdorff, 1978), and it has to be admitted that beyond the major factors there is little agreement, and an infinite regress of splitting factors up into finer and more useless subfactors (Carroll, 1993). Our concern here is only with g , and the frequently voiced objection in this connection that it is impossible to derive a unique value for g , so that we can give a *unique* score for intelligence to any given person. There are many ways of combining different tests; is there a "best" way? Jensen and Weng (in press) have attacked the problem with some success. After a careful statistical evaluation, they concluded that the two most common methods for extracting g , namely, the first unrotated factor in either a principal components analysis or in a principal factor analysis, are theoretically unsuitable methods. These and other methods differed quite markedly in overall accuracy of estimation of g in artificially created matrices. Even the popular Schmidt-Leiman procedure, although doing a fairly good job, was never quite accurate. The best method to use turned out to be Joreskog's LISREL (Linear Structural Relations) method, which is a maximum likelihood and "confirmatory" factor analysis method. It was also found that g was in fact extremely robust across 10 different methods of extraction, the mean of congruence coefficients amounting to .99+.

Another objection frequently voiced, and answered by rigorous statistical argument by Jensen and Weng, relates to the composition of the test battery, and the likelihood that its g would correspond with the true g . Given the absence of intentional selection of tests measuring one of the major primaries, this correspondence depends on the number of tests randomly selected; for the 11 tests of the Wechsler Intelligence Scales it would be slightly over .90. In fact, Wilks's theorem tells us that the choice of any particular method of extracting g becomes less and less important as the number of variables that are entered into the factor analysis increases. Thus the accuracy of measurement of g is determined largely by the number of tests used, and the number of persons tested. It approximates unity when both are reasonably large.

It would not be fair to say that everyone agrees with the postulation of g as a single major factor defining psychometric intelligence. Horn (1968, 1977, 1985) has consistently opposed this theory, and demanded what is in fact a 4-factor theory. His argument is fatally weakened by a disregard for studies using confirmatory factor analysis and multidimensional scaling; he might have had a point in the 1970s, but more recent work does not bear out his contentions. Neither does he seem to be willing to consider the capacity for achievement argument, which clearly puts g_f in a causal relation to g_c , even if we must agree that measures of g_f necessarily contain some earlier learning. On another point, Horn fails to consider the accuracy of Spearman's prediction as to what sort of test would have high g saturations. Finally, his arguments cannot explain why very different theories (e.g., Spearman's and Piaget's) show such high agreement psychometrically (Carroll, Kohlberg, & De Vries, 1984;

Eysenck, 1979; Humphreys & Parsons, 1979). Clearly one investigator's g is very much like another's.

The argument concerning g can of course be pursued at greater length (Eysenck, 1992), but there would be little point in doing so here. Let us merely note two things. In the first place, statistical argument, particularly factor analytic contentions, cannot in the nature of things resolve the conflict to everybody's satisfaction. Taxonomy always has a subjective side (Sokal & Sneath, 1963), and only a *causal* approach can resolve the issue. This demands testable theoretical statements and experimental testing. To this we will turn in the next section.

The other point relates to our conception of intelligence, and the use of the term. Spearman's g essentially defines something called psychometric intelligence, but the term *intelligence* is also used in two other senses, and discussion demands clarification of definition. Figure 1 shows the three terms in question and their relation. *Biological intelligence* refers to the basic anatomical, physiological, and hormonal properties and functions of the brain which underlie all forms of cognition, and mediate individual differences in cognitive ability. *Psychometric intelligence*, as discussed in this section, refers to g as measured by IQ tests. *Social or practical intelligence* (Sternberg, 1985; Sternberg & Wagner, 1986) refers to the successful application of IQ to the events of a person's life. These different meanings of the term are often confused, but they should be carefully kept apart in order to avoid confusion.

Biological intelligence clearly determines in large part psychometric intelligence, but there are also familial, educational, and socioeconomic factors which influence it, to the extent that environmental factors are found to determine it. The usual tally of 70% genetic–30% environmental contributions very roughly shows the relative contributions involved. The proportions differ from place to place and from time to time because heritability is a population statistic (Eysenck, 1973a, 1979). Practical intelligence is partly based on psychome-

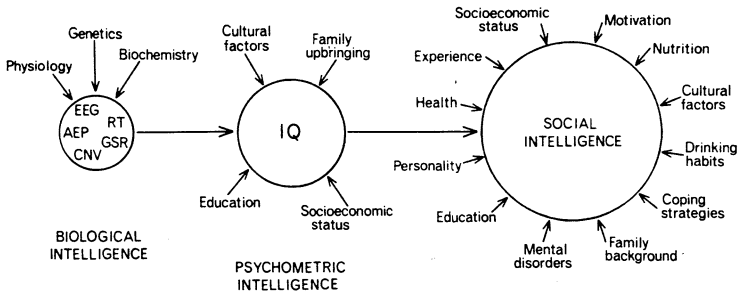


Figure 1. Three different types of "intelligence."

From "The Concept of 'Intelligence': Useful or Useless?", by H. J. Eysenck, *Intelligence*, 12, 1-16. Reprinted by permission.

tric intelligence, but even more influenced by many extraneous factors, as shown in Figure 1. It is such a complex concept that it has little scientific value, and we shall not discuss it any further (Eysenck, 1988).

The distinction between *biological* and *psychometric* intelligence does, however, give rise to theoretical debates which are fundamental to any theory of intelligence. The debate harks back to two protagonists of alternative ways of thinking about intelligence, namely Galton (1883, 1892) and Binet (1903, 1907). The essentials of this debate have been discussed in detail elsewhere (Eysenck, 1983, 1986b, 1986c). Essentially, Galton believed in a unitary concept of intelligence, and Binet in a multiplicity of functions or faculties, such as memory, imagery, imagination, attention, comprehension, suggestibility, persistence, and so on, all of which should be measured, with intelligence being merely the statistical average of relatively unconnected factors. (This is a rough-and-ready interpretation of Binet's theory; he never gave a detailed discussion of it, and changed his mind from time to time. See Tuddenham, 1962, for a more detailed account.) Galton may be regarded as the originator of *g*, Binet of the "primary abilities" movement; both now find a place in the hierarchical system pioneered by Burt (1909).

Galton laid stress on genetic factors in producing individual differences in intelligence, without disputing that environment also played a part. Binet, as an educational psychologist, was more interested in environmental manipulation, without denying the possibility of genetic determinants. Here, too, both were essentially right, but again Galton drew attention to the more important of two alternatives.

Finally, Galton, clearly preferring the concept of biological intelligence, suggested measurement by means of simple, biologically determined tests. He and Cattell favored simple tests of reaction time, sensory discrimination, time estimation, and so on. Binet preferred complex mental tests of a more cognitive character, and as we know he won the argument, and IQ tests have ever since been faithful copies of his original measure. Jensen (1982a, 1982b, 1986b) and Eysenck (1987b) have discussed the influence of the *Zeitgeist* which caused psychologists to accept very bad experiments (e.g., Wissler, 1901) as proof that Galton was wrong. The point to remember is that choice of test is intimately connected with *choice of theory*. Galton accepted a capacity theory, leading to attempted measurement of *biological intelligence* by means of simple *physiological measures*. Binet accepted a theory laying more stress on achievement and knowledge, and emerged with measures which owed a lot to *learning*. American theories have tended to *identify* IQ with acquired knowledge and to reject in large measure genetic and physiological factors (Sternberg, 1982; Wolman, 1985). The ranks are not quite so clearly defined, but the contrast is clear. Recent research has revived the Galtonian approach, and any theory of intelligence as *capacity* must accommodate this recent work.

THEORIES AT THE CAUSAL LEVEL: SPEED OF INFORMATION PROCESSING

The original work of Spearman (1904) and Burt (1909) actually followed in large measure Galton's precepts, with positive consequences. Spearman found tests of sensory discrimination a good measure of intelligence, and later work usually found positive correlations, sometimes unexpectedly high (e.g., Buktenica, 1971; Raz, Willerman, & Yama, 1987; Stankov & Horn, 1980). The Raz et al. work is undoubtedly the best controlled of these studies. The authors found correlations of -0.42 and -0.54 between discrimination thresholds and an IQ test in college students, that is, with seriously restricted range of ability. Correction for restricted range suggests a "true" correlation of about 0.60; this is a remarkable vindication of Spearman's original hypothesis. (Note that "discrimination" is one of the senses in which the Romans used the word *intelligentia*.)

Burt (1909) also found sensory discrimination tests correlated with intelligence; motor tests and sensorimotor tests also did well. An early version of an IT (inspection time) apparatus test gave particularly good correlation with scholastic intelligence. Given this very respectable performance of Galton-type measures, one cannot help but wonder why they were abandoned by Spearman and Burt in favor of Binet-type tests. Perhaps the need for apparatus use, laboratory equipment, and individual testing caused this sea change.

It is of course clear that reaction-time and inspection-time measures are essentially tests of *discrimination*, although not purely sensory; there is always a (small) cognitive content. It might also be noted that discrimination is the major ingredient in the only type of tests used on neonates to predict later IQ (Fagan, 1984; Fagan & McGrath, 1981). In those tests what is measured is the neonate's preference for novel stimuli; this presupposes discrimination between novel and repeated stimuli, as well as memory for previous exposure. But memory has always been found to show low correlation with IQ, so that the major role in the prediction would seem to be borne by discrimination. The importance of discrimination in defining the nature of intelligence fits in well with the speed-error theory to be described later in this chapter.

It is only relatively recently that attention has returned to Galton-type tests, particularly reaction-time and inspection-time tests (Carlson & Widaman, 1987; Eysenck, 1982, 1986b; Jensen, 1982a, 1982b, 1985, 1986a). The major findings are that RT measures (both decision time and movement time), correlate negatively with IQ, poorly for simple RT, more highly the more complex the demands of the task (Frearson & Eysenck, 1986); that *variability* of performance correlates even more highly (negatively) with IQ; and finally that the speed or power nature of the IQ test is irrelevant to the size of the correlation. Zero-order correlations of 0.60 and above have been found for RT-IQ (e.g., Frearson & Eysenck, 1986), and multiple Rs for combinations of RT measures

of up to 0.70 and beyond. For ITs similarly mean correlations of around 0.50 are reported. Correlations are lower for both IT and RT when range of IQ is restricted, and higher when retardates are part of the sample. Altogether a mixture of relatively simple RT-IT measures, easy enough for retardates to do without error when given enough time, can correlate with a good IQ test as highly as one good IQ test with another. This is a remarkable finding, as is the finding that the correlation of a Wechsler subtest with g is directly proportional to the correlation of that test with an RT measure ($r = 0.83$, Hemmelgarn & Kehle, 1984), suggesting that the RT test was a direct measure of g .

Can we assert that these data prove that a factor of mental speed underlies g as Eysenck (1953, 1967) suggested? Detterman (1987) has argued that we cannot, and of course he is right; no such inductive argument would be admissible. But the argument is a different one. I would suggest that the facts regarding RT and IT discovered recently *support* deductions from a theory suggesting that some form of mental speed underlies g —in other words, it is not that a theory is suggested by serendipitous findings, but that the findings are predicted on the basis of a theory, and to that extent support it. There may of course be alternative interpretations, and new theories may be formulated to account for the facts. But insofar as the facts are in line with prediction, they support the theory. Detterman also argues that RT experiments are more complex than they seem, and again he is right. The question that arises is whether the complexities he lists are likely to *cause* the observed correlations, and the answer seems to be rather that they would *lower* them, but this point is of course subject to experiment. What is the effect of motivation, and changes in motivation brought about by knowledge of results, say, or rewards and punishments, on RT-IT correlations? One of the major functions of a theory is to generate research, and clearly this theory has done precisely that.

At this point, we may notice certain remarkable similarities between the development of the concept of G (gravitation) and g . Note that both arose from simple observations. *Gravitas* (weight, heaviness) is the product of widely observed phenomena (bodies unsupported falling to the ground), whereas *intelligentia* equally represents widely observed cognitive phenomena, and individual differences. Newton and Spearman, respectively, reported certain unifying hypotheses, expressed in mathematical form, leading to the postulation of scientifically meaningful concepts, G and g . Neither theoretical and mathematical development was rigorous; calculus only became mathematically rigorous 150 years after Newton when Cauchy published his *Cours d'Analyse*. Both theories were beset by anomalies, which only "ordinary science" (Kuhn, 1970) could explain. Both theories were heavily criticized in a nationalistic spirit—Newton's by the French, Spearman's by the Americans. Both suffered from an inability to explain certain phenomena: Newton was plagued by the three-body problem, Spearman by the existence of "primary" or group factors. Neither author succeeded in finding a causal theory to explain his descriptive findings,

seeking recourse in such insubstantial notions as “action at a distance” (Newton), and “mental energy” (Spearman).

Lack of any such theory to explain the phenomena of intelligence testing led the writer to take up some vague theories and notions dating back to Thorndike et al. (1926). It was argued (Eysenck, 1953, 1967, 1982, 1987b) that the usual method of taking total scores on an IQ test as a measure of intelligence was inappropriate, because different individuals could obtain identical scores with quite different item solutions. It was suggested that the proper unit of analysis was the item, which could be correctly solved, incorrectly solved, or abandoned. In addition, it seemed desirable to *time* each solution. The resulting model suggested that IQ measurement depended on three theoretically independent factors: mental speed, persistence or continuance, and error production. A lengthy period of empirical and theoretical work followed, in which Furneaux (1952, 1961) took a prominent part. Later on, White (1973, 1982) elaborated some statistical solutions to some of the problems which arose. Berger (1976, 1982), Brierley (1961), and others also reported data relevant to the theory.

Furneaux (1961) was responsible for elaborating a model which is fundamental for any understanding of intellectual functioning; his theory can best be stated in his own words:

The brain structure of any individual, P, includes a set of p^N neural elements which participate in problem-solving activities. It is not necessary at this stage to adopt any particular view as to the nature of these elements, which might be either single neurones or much more complex structures. The solution of a particular problem, h, of difficulty, D, involves bringing into association a particular set, D^N_h , of these elements, interconnected in some precise order. (The terms ‘bringing into association’ and ‘interconnected’ should not necessarily be interpreted literally after the manner of, say, an electrical circuit. For example, the almost simultaneous firing of two otherwise independent units could constitute one method of bringing them into association, provided some device existed which could detect the simultaneity, while the exact order of firing might represent the mode of interconnection.) When problem h is first presented single elements are first selected, at random, from the total pool p^N and examined to see whether any one of them, alone, constitutes the required solution. A device must be postulated which carries out this examination—it must bring together the neural representations of the perceptual material embodying the problem, the rules according to which the problem has to be solved, and the particular organization of elements whose validity as a solution has to be examined. It must give rise to some sort of signal, which in the case of an acceptable organization will terminate the search process and will initiate the translation of the accepted neural organization into the activity which specifies the solution in behaviour terms. Alternatively, if the organization under examination proves to be unacceptable, a signal must result which will lead to the continuation of the search process. It

will be useful to refer to this hypothetical device under the name of 'the comparator.'

If $D \neq 1$, the comparator, will reject each of the p^N trial solutions involving only a single element, which, when correctly interconnected, might constitute a valid solution. Suppose $D \neq 2$, then all possible organizations of the p^N elements taken two at a time will also be examined and rejected, after which we can imagine that the search will continue among sets of three, four, five, etc. If $D = r$, then the comparator will reject in turn all the organizations involving from 1 to $(r - 1)$

elements, so that there will be a time $r \sum_{r-1}^1 E$ sec within which a solution cannot occur, where:

r = the time required for completing a single elementary operation within the search process.

$r \sum_{r-1}^1 E$ = the number of elementary operations involved in the search process up to the level of complexity $(r-1)$.

Similarly, after a time $r \sum_r^1$ sec all possible organizations embodying r ele-

ments will have been examined, so that correct solutions to problems of difficulty r will always arise within the period defined by the two limiting times:

$$r \sum_{r-1}^1 E \text{ and } r \sum_r^1 E$$

In terms of such a hypothesis, therefore, V_{Tr} , is in no sense a function of error of measurement but results mainly from the range of times required to set up all possible modes of neural organization at a particular level of complexity. It is perhaps worth noting, in passing, that within the framework of such a hypothesis, error would be accounted for by positing that during the search process organizations arise at levels of complexity $v - \delta_1, v - \delta_2, \dots$, etc., which satisfy most, but not all, of the requirements of a true solution to a problem of difficulty r . If the comparator has characteristics analogous to those of "band-width" in electrical and mechanical discriminators, i.e., if its discriminating powers are such that neural organizations which closely resemble the organization representing a correct solution may be accepted as the required organization, then the possibility of error arises. The frequency of error, thus conceived, will be a function of the band-width of the comparator, and since the number of "nearly-correct" organizations will increase as D is increased, the likelihood of error will increase with D . This probability is clearly dependent on the exact nature of the search process. Finally, continuance is easily defined in terms of such a "search" hypothesis; it is a measure of the length of time during which, following the initiation of search, the comparator remains "set" for a particular problem. (pp. 185-186)

The most recent test of the model against the Wechsler and RT measures (Frearson, Eysenck, & Barrett, 1990) indicates its present position. We carefully summarized both the positive and the negative conclusions to be drawn from our studies.

The Furneaux model approaches individual differences in problem solving in a new way. Both the properties of items and Ss are reinterpreted. Items are considered in terms of their difficulty, based not on the probabilistic properties of pass/fail (as in classic psychometrics and item-response theory) but rather in terms of difficulty based on time to solve. The relationship between this 'time based' difficulty and the difficulties of classical psychometrics and item response theory remains unexplored, but the meaningfulness of time-based difficulties (i.e. that relative item difficulty is dependent solely upon the item's content not upon properties of the sample's individuals) has been demonstrated.

The classification of Ss by their speed, accuracy and continuance has also been shown to be reliable and apparently valid. However, the strict adherence to the Furneaux method for the computation of S speed and continuance parameters which leads to the failure to assign a full set of parameters to many Ss (and hence has led to the premature halting of attempts at replication) has been shown to be unnecessary. Parameters for continuance can be computed directly from the times taken to abandon items and such parameters out-perform parameters calculated with a closer adherence to the letter of the Furneaux method. Similarly, S accuracy parameters can be computed with no regard to the Furneaux model (indeed it is hard to see how the Furneaux method does calculate accuracy parameters) and appear to perform well.

It is in the calculation of speed parameters that the Furneaux model shows distinct advantages over more crude measures. To investigate individual differences in speed it is clearly imperative to account for differences in item characteristics. A simple summation of the time to complete a set of items can never show the relationship between speed and general ability as it confounds two opposite tendencies: the tendency for bright Ss to rapidly complete items and to succeed at difficult and time-consuming problems. Attempts to measure speed have therefore concentrated on giving Ss items only of such low difficulty that all Ss will succeed on them, measuring only the first of the tendencies. The resulting speed measures applying as they do only to items of a trivial difficulty are of limited validity (and the difficulties in measuring solution times are magnified by virtue of their shortness). The Furneaux method addresses the problem of measuring speed by determining for items a time-based difficulty measure. The use of these difficulties means that Ss' speeds can be calculated from any set of items of a full range of difficulty. It is this facility that allows the Furneaux method to produce speed measures for subjects that are reliable and apparently valid.

The nature of this speed characteristic of individuals is indicated by the results from comparing Furneaux speed and the WAIS-R scores and decision time data. Significant correlations between Furneaux speed and the unspeeded subtests of the WAIS-R, particularly vocabulary, show that Furneaux speed is not simply a facility for working against the clock but is rather a trait for rapid manipulation

of data. The low correlation between speed and decision time parameters shows Furneaux speed not to be a simple measure of ability for rapid response but rather an aspect of more cognitive traits.

Theoretically the poor correlation between the Furneaux speed measure and decision time parameters is counter indicative of a simple hierarchical model where fast decision times lead to fast cognitive speed and hence a high IQ, but rather points to a model of a high IQ dispositional set where responsivity to outside stimuli interacts with internal properties of the nervous system leading to the parallel development of fast decision times, cognitive speed and IQ.

The practical applications of the Furneaux method are important. It has long been felt that the time taken for a S to produce a response was as informative as whether it was right or wrong. However, the difficulty in accurately timing each response has made use of such data prohibitive. With the introduction of automatic ability testing by micro-computer, the collection of response times becomes trivial. The Furneaux method allows use to be made of these extra data. The results of the multiple regressions using Furneaux speed along with accuracy and persistence parameters to predict WAIS-R subtest scores as well as verbal, performance and full-scale IQ shows that the use of response times allows a short test to perform more like a battery of tests than a single one. (Typically any one test will correlate highly with some tests with which it shares common item-specific factors and very badly with others. The generally high multiple R obtained with the three parameters of the Furneaux test for all the WAIS-R subtests show its greater independence from the effects of item specific factors.)

The supporting evidence for the details of the Furneaux model is weak, indeed it is hard to see what form such evidence could take. However, the function of models of such wide scope is more to be of use in stimulating research than to be literally correct. The Furneaux model highlights an alternative approach to individual difference research distinct from both psychometrics and item response theory. Such an approach appears to give results of acceptable reliability and validity and with changes in techniques of practical testing the overheads involved in collecting the extra data required become less of an obstacle. (Frearson, Eysenck, & Barrett, 1990, pp. 256–257)

This approach is in many ways similar to a theory put forward by Campbell (1960) at the same time as Furneaux's final presentation of his theory. Simonton (1988) has taken up the Campbell theory and elaborated it recently. Campbell proposes a "blind-variation-and-selective-relation process" (p. 380) as fundamental to all inductive achievements, and to all genuine increases in knowledge, as well as to creative thinking. This would thus furnish a basis for Spearman's noegenetic laws (although neither Campbell nor Simonton put it like that). Campbell argues that a similar view has been put forward by philosophers and scientists, such as Bain, Sourian, Mach, and Poincaré. Campbell does not formalize his system as Furneaux has done; this makes it impossible to test. But the similarities are noteworthy, as is the simultaneous emergence of the two theories. Their main import—and it is an important

one—is the absence of a homunculus in the brain doing all the theorist's work for him!

We now have a theory of cognitive functioning, and we have a theory of individual differences in mental speed which, taken together, may have some value in suggesting a *causal* basis for differences in *g*. It clearly cannot be asserted that this is a truly causal relation. In fact it will be suggested in our next section that *speed* is a secondary factor only, the true causal factor being the error rate in information processing. But staying with speed for a moment, we may ask the somewhat naive question of whether speed of sensory nerve conduction might be the physiological basis of the observed differences in RT, IT, and mental speed generally. Barrett, Daum, and Eysenck (1990) have reported an experiment in which orthodromic electrical stimulation of the median nerve at the site of the third digit in each hand was used to elicit average sensory nerve action potentials. The outcome was that no correlation appeared between *velocity* of sensory conduction and IQ, but that *variability* of sensory conduction did so ($r = 0.44$). Removal of outliers increased this correlation to 0.68. Vernon and Mori (1989) have reported a correlation between intelligence and nerve conduction velocity of 0.42, and between RT and nerve conduction velocity of -0.28. The true correlation thus remains in doubt, although a recent publication (Barrett & Eysenck, 1993) and as yet unpublished study by Vernon would seem to argue against the theory. What is apparent is that although there is support for a mental speed hypothesis, there are two findings which argue against it, or at least constitute anomalies which cannot be interpreted in these terms. One involves the variability of sensory conduction and its correlation with IQ, the other the variability of RT measures and its correlation with IQ. Some such correlation might be predicted from the obvious connection between RT speed and variability; high speed demands low variability, as high variability implies many low-speed reactions. But what is found is usually a *higher* correlation for variability than speed, and this cannot be so explained. An attempt to do so will be given in the next section.

CAUSAL THEORIES AT THE BIOLOGICAL LEVEL

Reaction times and inspection times are relevant to Galton's conceptions, but they are not strictly physiological measurements of brain activity. Recent years have seen a burgeoning of approaches to this topic, using EEG, evoked potential, and contingent negative variation measurements for the purpose (Deary & Caryl, 1993; Eysenck, 1986a, 1986b, 1993a; Eysenck & Barrett, 1985; Vernon, 1993). The results leave little doubt about the fact that quite high correlations can be obtained between particular AEPs and IQ. This is not the place to recapitulate what is in fact an enormous amount of evidence. We shall be con-

cerned rather with the question of the light selected studies throw on the theory of intelligence.

Much work has shown that the *latency* of AEPs is negatively correlated with IQ, very much like RT and IT; the slower reactions, the lower the IQ (Callaway, 1975; Deary & Caryl, 1993). Amplitude also usually showed a positive correlation with IQ. More important theoretically is the finding that *variability* of AEP gives more consistent and higher correlations with IQ than either latency or amplitude. Earlier work is summarized by Deary and Caryl (1993), but for our purposes the most important study is one reported by D. E. Hendrickson (1982) and A. E. Hendrickson (1982). Measures of complexity of the wave form and of variability were taken on 219 schoolchildren, with a mean WAIS score of 108. The measures of complexity and variability were of course quite highly correlated with each other, and both correlated with IQ (0.72 and -0.72, respectively); the composite score correlated with IQ 0.83. Complexity was measured in terms of the total length of the trace ("string test" because in the early work a string was laid along the trace, and then straightened and its length determined). Variability was determined by adding the variances of the response to the 100 tones which constituted the stimuli on each of the 2 millisecond recordings.

Is the correlation between *g* and AEP a proportional measure, as theory would demand? Eysenck and Barrett (1985) carried out the following calculation.

If the general factor obtained from the intercorrelations between all the subtests of the Wechsler is our best index of intelligence and if the AEP Composite measure represents a good measure of intelligence, then we would expect the *factor loadings on the 11 WAIS subtests and the correlations of the subtests with the AEP Composite measure to be proportional*. Table 1 shows the actual data, giving both factor loadings and correlations with the AEP Composite measure. In view of the fact that the reliabilities of the different WAIS subtests are not identical, we give both the uncorrected (raw) correlations and the correlations corrected for attenuation. It will be seen that as far as the correlation between factor loadings and composite measure is concerned, the correction makes little difference; the Spearman rho is .95 for the uncorrected values and .93 for the corrected values. Proportionality, therefore, is almost perfect and strongly supports the view that the EP is a true measure of intelligence.

Why do we find such a close relationship between IQ and variability? A. E. Hendrickson (1982) has suggested that the reason is the occurrence of errors in the transmission of information through the cortex, and that such errors produce variability. He provisionally located the source of the errors in the synapse, and framed the argument in terms of pulse trains forming a chain, with the final pulse train being "recognized" in order to initiate action. His argument is as follows:

Table 1. Correspondence Between Factor Loadings of 11 WAIS Tests and their Correlations with Composite AEP Measure

WAIS subtests	Correlations with			
	Composite AEP measure		Factor loadings	
	Uncorrected ^a	Corrected ^b	Uncorrected ^a	Corrected ^b
Information	-.68	-.71	.78	.82
Comprehension	-.59	-.66	.73	.82
Arithmetic	-.65	-.73	.78	.88
Similarities	-.71	-.76	.82	.88
Digit Span	-.59	-.70	.68	.81
Vocabulary	-.68	-.70	.79	.81
Digit Symbol	-.35	-.36	.50	.52
Picture Comprehension	-.57	-.63	.68	.75
Block Design	-.54	-.58	.71	.77
Picture Arrangement	-.46	-.57	.58	.71
Object Assembly	-.44	-.55	.58	.72
The AEP Composite Measure	—	—	-.77	—

^a The Spearman rho between the *Uncorrected* correlations and loadings = .95

^b The Spearman rho between the *Corrected* correlations and loadings = .93.

From "Psychophysiology and the Measurement of Intelligence," by H. J. Eysenck & P. Barrett, 1985, *Methodological and Statistical Advances in the Study of Individual Differences* (Eds. C. R. Reynolds & V. Willson), New York: Plenum. Reprinted by permission.

Let us call the characteristic probability of correct recognition for an individual 'R'. The converse probability of recognition failure is therefore 1 - R. R is to be thought of as the probability that just one of the synaptic recognitions will succeed, as opposed to the probability that the entire chain will be recognized correctly.

If we assume that each synapse has the same value of R, and that the probabilities are independent, we specify the probability that a chain of 'N' events will succeed as:

$$R^N$$

that is, R raised to the Nth power. For example, if R was .90, and N was 3, the probability of the 3-chain sequence would be .729.

Let us reverse the problem somewhat, and examine the characteristic length of a logical chain that will succeed before breakdown occurs, given the value of R. This is called the expected average value of N, written as E(N). The formula turns out to be:

$$E(N) = \frac{1}{1 - R}$$

There are some very interesting behavioural consequences implied by the above formula. We should note that the value of R is likely to be quite high in absolute terms, and it is interesting to see the effect that small differences in the value will make.

Suppose that we have three individuals, which we will call 'A', 'B', and 'C', each having a different characteristic value of R . Individual 'A' has an R value of .9900, individual 'B' a value of .9990, and individual 'C' the highest value of all, .9999.

Substituting the values of R into the above formula, we can work out that the average expected values of N , $E(N)$, are 100 for individual 'A', 1,000 for individual 'B', and a full 10,000 for individual 'C'!

Converting the above expected values into time, we use the value of 230 ms (our assumed average human pulse train length) for each of the N units. Individual 'A' has an expected mean time before failure (MTBF) of 23s. Individual 'B' has a MTBF of 230s, or just under 4 min. Individual 'C' has a MTBF of 2300s, or about 38 min.

For many kinds of everyday activity, we humans do not require very long thought chains. If a particular thought chain was one that required, say, only ten steps for completion, there would be very little absolute difference between the values of R to the N in the three individuals mentioned above. Individual 'A', with the R of .9900, would have a probability of completing a ten-link chain of .904. Individual 'B', with an R value of .9990, has a final probability of .99 of completing the chain, and individual 'C' ends up with a .999 probability.

If we had some sort of IQ test which consisted of large numbers of easy items, each of which could be solved using programmes of ten links or less, and these programmes were known to our three individuals, it would take a fairly large number of test items to be administered before we could reliably identify individual 'A' from the others. A 100-item test, with each item requiring ten pulse train links, and a time limit placed on each item, should result with 'A' having a score of 90 on the test, 'B' with a score of 99, and 'C' with a score of 100. We would not be very confident on the basis of these results that 'B' and 'C' were really different, and not entirely certain (taking into account normal sampling variability) that 'C' was really better than 'B'.

Now let us step up the difficulty by a factor of 100, and set a test consisting of items with 1,000 links each. Each item requires just under 4 min of thinking to arrive at the correct answer. Individual 'A' breaks down on this test completely. He has a probability of solving any single item of only .00004—lucky guesses excluded. Individual 'B' has a respectable probability of solving any item of .368. Individual 'C' still finds the items fairly easy, with a probability of .904 of getting the answer to an item in the allotted time. If we were to administer a timed test with ten such items to these three individuals, we would expect scores of 0, 4, and 9, which spaces the individuals out nicely. We would therefore be fairly confident that we have distinguished between them with a small number of items (noting, however, that the ten items take longer to administer, as a practical point).

Our model mirrors the commonplace observation that people of very high

intelligence (as inferred from their jobs or accomplishments) are not vastly superior in everyday tasks to people of ordinary intelligence. The model also fits in with the feeling of many teachers that it is virtually impossible to teach very difficult tasks to some people, no matter how long one persists (for example, higher mathematics to ESN children). (A. E. Hendrickson, 1982, pp. 184–185)

This “error probability” model is not incompatible with our “speed of information processing” model. I have argued (Eysenck, 1987b) that the error model implicates the speed differences between high and low IQ groups, and has precedence because it explains the correlations between IQ and variability (RT; AEP) which the speed model leaves unexplained. Consider Figure 2, which is a greatly oversimplified model of information processing in the brain. A “sender” receives a sensory input, say a tone as in the Hendrickson experiment. This message is relayed over a series of hundreds of neurons (represented by just 8 in the figure) which pass the message on to a comparator which has to decide about the *nature* of the message before initiating action. If all incoming messages agree on the nature of the stimulus, the comparator can fire straightaway (no error paradigm). But let us say that in some of the neurons or synapses an error occurs; this results in ambiguity of the message at the comparator level, and further input needs to be awaited before the message is clear and acceptable. As messages are not sent out simultaneously but over a (very short) period of time, each needs varying lengths of time to arrive at the comparator, errors will cost time, and this time loss means a less speedy reaction. Errors are of course only probabilistic; they may or may not occur even in an individual with high error probability. Hence on the occasional error-free occasion, even the dullest retardate may give an RT as short as a high-IQ subject; it is this which produces the correlation between variability and IQ. In other words, speed differences are derivations of error probability differences. They

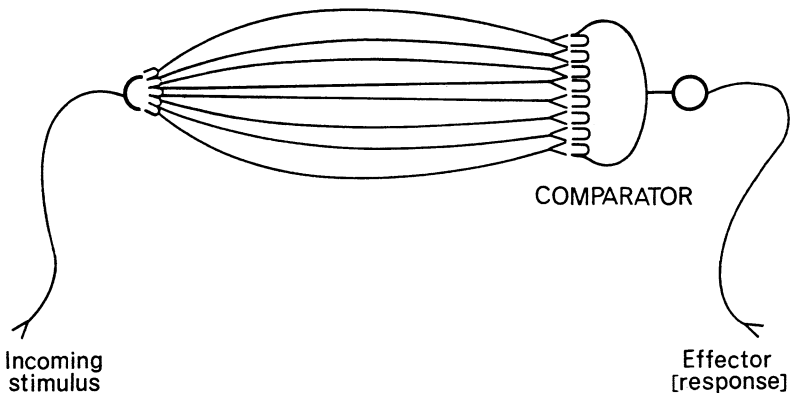


Figure 2. Diagrammatic picture of cortical comparator.

are only likely to give correlations with IQ in very easy tasks; in tasks lasting over a second, and admitting of executive error, the picture becomes more complex, and errors, omissions, "giving up," and differential strategies may play a part. Individual differences in IQ tasks would still be *caused* by transmission error probability, but would not necessarily act through speed differences as measured. It seems that log times to solution depend on the difficulty level of the problem. These soon become too long to be tolerable, and hence the subject gives up, or produces a premature wrong answer (Eysenck, 1953). How the individual reacts depends on his or her personality; that is, noncognitive factors designated "continuance" and "error" in the Furneaux model. Extraversion, impulsiveness, and the like are clearly involved.

The theory suggested here would explain not only the importance of *g* in cognitive tasks in terms of a single fundamental psychological-physiological process, it would also serve to explain many other findings hitherto rather mysterious. Why are correlations of mental tests with each other and with cognitive variables highest for low-IQ groups (Detterman & Daniel, 1989)? Detterman argues that possibly central mental processes, if deficient, would limit the efficiency of all other processes in the system.

Because of the deficit in the central process, the entire system is brought to a uniform low level of operation. So all processes in subjects with deficits tend to operate at the same uniform level. However, subjects without deficits show much more variability across processes because they do not have deficits in important central processes. This causes high correlations among mental measures in low IQ subjects and low correlations in high IQ subjects. (p. 358)

If we identify error propensity with this central process, we obtain a testable version of this theory. It gives *causal* backing to the *descriptive* theory of *g* as a general factor of intelligence.

Any summary of the evidence concerning AEP relationship with IQ (Deary & Caryl, 1993) must note the *variability* of results reports. This is due in part to different choices of IQ measures, different choices of electrode placement, differences in EEG methodology, differential samples (restriction of range), and similar experimental factors. Barrett and Eysenck (in press) have recently shown that the size of the correlations obtained depended crucially on the amplitude of the P180 component. High-amplitude subjects gave high correlations; low amplitude subjects low correlations. This relation was obtained in three different samples, and was also found in an inspection time paradigm. In other words, subjects are *heterogeneous* as far as biological variability is concerned, and may thus give rise to contradictory results if no measure of this factor is obtained. The psychological meaning of P180 amplitude in this respect is discussed by Barrett and Eysenck (in press).

CREATIVITY AND OTHER NON-G FACTORS

We have deliberately concentrated on the elaboration of a theory to account for *g*, leaving out of account non-*g* factors or "primaries." Some of these non-*g* differences may be due to cerebral specialization (Kaufman, 1979) or other factors; there is little evidence on which to base a theory. But IQ tests are often accused of neglecting important factors like creativity and originality, and it is the intention in this section to suggest a model of *creativity* which fits into the framework outlined in previous sections.

Originality and creativity are often taken to be cognitive traits, that is, aspects or parts of intelligence. Thus intelligence tests are often divided into measures of *convergent* and *divergent* thinking, with the latter being more closely associated with originality and creativity. An alternative view, which is taken in this section, is that originality and creativity are not by themselves aspects of intelligence, but rather are traits of *personality*, that is, are noncognitive. Great achievement, on this account, would be due to a combination of high intelligence and the appropriate personality configuration. This view, which was originally put forward by Spearman (1927), requires empirical support, and such support has recently been forthcoming in a series of studies, most of which were conducted by British psychologists.

Some of this work took its origin from the widely held hypothesis that genius and madness may be closely allied. This common observation suggests that people who are highly original and creative may differ from the ordinary run of people by showing personality qualities often associated with schizophrenics and other psychotics. A number of genetic studies have indeed supported such a view. Heston (1966) studied offspring of schizophrenic mothers raised by foster parents, and found that although about half showed psychosocial disability, the remaining half were notably successful adults, possessing artistic talents and demonstrating imaginative adaptations to life to a degree not found in the control group. Karlsson (1968, 1970) in Iceland found that among relatives of schizophrenics there was a high incidence of individuals of great creative achievement. McNeil (1971) studied the occurrence of mental illness in highly creative adopted children and their biological parents, discovering that the mental illness rates in the adoptees and in their biological parents were positively and significantly related to the creativity level of the adoptees. Findings such as these clearly support speculations, such as those by Hammer and Zubin (1968) and by Jarvik and Chadwick (1973) to the effect that there is a common genetic basis for great creative potential and for psychopathological deviation.

Eysenck and Eysenck (1976) have suggested that psychoticism might be a causal factor in creativity. Assuming for the moment that the P scale does measure, at least to some extent, the essence of the continuum from normality to psychosis, and assuming for the moment that the hypothesis linking creativity

and originality with mental abnormality possesses some virtue, then we should be able to test this hypothesis in a variety of ways. It was first tested, in an unpublished study referred to by Eysenck and Eysenck (1976), by Kidner.

He administered several of the Wallach and Kogan (1965) tests of originality to male and female students, nurses, and teachers, and found significant relationships between originality and creativity, on the one hand, and high P scores on the other. He also found that *acceptance of culture*, that is, agreement with cultural norms, was negatively correlated with P, and also with creativity and originality.

Other studies more marginally relevant to the hypothesis under investigation are reported in the book by Eysenck and Eysenck (1976), but we will turn now rather to a more recent study by Woody and Claridge (1977) which is particularly impressive.

The subjects of their study were 100 university students at Oxford, both undergraduate and graduate. The students constituted a wide sampling of the various fields of specialization at the university. They chose students as their subjects because of evidence that creativity is significantly related to IQ up to about 120, but that it becomes independent of IQ above this level. The tests used by them were the EPQ (Eysenck & Eysenck, 1975) and the Wallach-Kogan Creativity Tests, somewhat modified and making up five different tasks (instances, pattern meanings, uses, similarities, and line meanings). Each task was evaluated in terms of two related variables: the number of unique responses produced by the subject, and the total number of responses produced by the subject.

The Pearson product-moment correlation coefficients between psychoticism and creativity scores for the five tests are as follows. P with total number of responses include: Instances = 0.32; Pattern Meanings = 0.37; Uses = 0.45; Similarities = 0.36; Line Meanings = 0.38. P with uniqueness scores: 0.61, 0.64, 0.66, 0.88, 0.65. It will be seen that all the correlations are positive and significant, and those with the uniqueness score (which is of course the more relevant of the two) are all between .6 and .7. These values are quite exceptionally high for correlations between what is supposed to be a cognitive measure, and a test of a personality trait, particularly when general intelligence has effectively been partialled out from the correlations through the selection of subjects. There were effectively no significant correlations between E and N, on the one hand and creativity on the other. It is interesting to note, however, that the L (Lie or dissimulation) score of the personality questionnaire, which up to a point is a measure of social conformity, showed throughout *negative* correlations with creativity scores, seven out of ten being statistically significant. L is known to correlate negatively with P (Eysenck & Eysenck, 1976).

Studies not using the P scale have come up with traits of creative persons not dissimilar to those characteristic of the high P scorer. Getzels and Jackson (1962) found that *divergers* were unconventional and independent of judgment

(see also Torrance, 1962). Hudson (1966, 1968) also noted the conformity of the converger, and the rebelliousness of and failure to *fit in* of the diverger.

It might be said in criticism of the studies reviewed so far that they deal with psychological tests of creativity and originality in normal and not very distinguished people, and that what is normally understood by originality and creativity demands something more than that. The objection is a reasonable one, although it should not be taken to infirm the remarkable success achieved by Woody and Claridge's empirical testing of the hypothesis linking P and creativity. The only study of what most lay people would consider genuine creativity has been reported by Götz and Götz (1979a, 1979b).

In the study under review, Götz and Götz administered the EPQ to 337 professional artists living in West Germany, of whom 147 male and 110 female artists returned the questionnaire: Their mean age was 47 years. One outstanding result of this work was that male artists were significantly more introverted and significantly more neurotic than nonartists, whereas for females there was no difference on either of these dimensions. As the authors suggest, it is perhaps true that in our Western world it is mainly women with average or higher scores on extraversion who have the courage to become artists, although the more introverted and possibly more artistically gifted women do not dare to enter the precarious career of the artist.

We must now turn to scores on psychoticism. Here the results are very clear; male artists have much higher P scores than male nonartists, and female artists have much higher P scores than female nonartists. As Götz and Götz point out, these results suggest that certainly many artists may be more tough-minded than nonartists: "Some traits mentioned by Eysenck & Eysenck may also be typical for artists, as for instance they are often solitary, troublesome and aggressive, and they like odd and unusual things" (p. 332).

The work of Götz and Götz thus offers important support for the results of Woody and Claridge, and the other authors cited above, in that this more recent work uses actual artistic achievement as a criterion for the measurement of creativity and originality. In doing so, they give credence to the validity of divergent thinking tests as measures of creativity and originality, and the fact that both in the artistic and in the nonartistic population studied by other investigators significant correlations are found between psychoticism and creativity and originality very much strengthens the hypothetical link between the personality trait and the behavioral pattern. We may thus be justified in concluding that originality and creativity are the outcome of certain personality traits, rather than being cognitive variables or abilities. This is an important conclusion which is somewhat in contrast with assumptions usually made in this field.

The Götz and Götz study is the only one which actually used the psychoticism scale, but other studies implicated traits in creative people which are clearly part of the P syndrome. Thus work of the Institute for Personality

Assessment and Research at Berkeley, under the direction of MacKinnon (1962), was concerned with creativity in architects, writers, and mathematicians. As described by MacKinnon et al. (1961) and Barron (1969), creative people showed traits of individualism and independence, lack of social conformity, unconventionality, and lack of suggestibility (Crutchfield, 1962). They were also below par in sociability and self-control. Responses on tests like word association were odd and unusual, almost like those of schizophrenics.

Most important, the creatives studied by the I.P.A.R. group consistently showed greater psychopathology on the Minnesota Multiphasic Personality Inventory depression, hypochondriasis, hysteria, psychopathy, and paranoia scales than did the controls. Lytton (1971) concludes that: "It is difficult...to deny that there is more than a chance association between psychiatric difficulties and creative powers" (p. 63). This psychopathology is countered, however, by greater ego strength, as also shown on the MMPI scales.

Of course, it should not be assumed that this personality trait of P, even when found in conjunction with high N and low E, can by itself produce an original work of consequence. A certain reasonably high amount of intelligence and/or artistic ability is obviously required in order to enable a person possessing high creativity and originality to produce anything worthwhile. It is obviously important to separate the successful use of personality traits—such as those mediating creativity—and the unsuccessful use, degenerating into mere oddity and possibly psychotic deterioration.

We would thus conclude that creativity *requires* high *g*, but is in itself a personality trait related to P. Can we specify more closely how it might be connected with our model of intelligence? We have postulated that cognition follows the "blind-variation-and-selection" rule, but of course for most purposes some extensions and restrictions of this rule are required. Selection can only take place between existing elements, and these have to be acquired by learning. Newton would not have produced his famous formula if there had not been in his mind the results of Galileo's and Kepler's researches, for instance. Again, not all known elements enter into the selection process. There were many facts known to Newton which did not enter into the "variation-and-selection" process, such as the name of his mother, or the assumed relation between Zeus and Athenae. Thus the process cannot be, and is not, entirely blind; there is implied a *restriction by relevance*. This point needs to be made, although it plays no part in the Furneaux or Campbell expositions.

One fundamental characteristic of schizophrenic thinking is very relevant to this theory. As Cameron (1939; Cameron & Margaret, 1951) and Payne and Hewlett (1960) have shown experimentally, *overinclusion* is very characteristic of schizophrenic thinking, that is, the tendency to go beyond the bounds of what is traditionally considered relevant in one's problem-solving behavior and include apparently irrelevant items. This will usually be counterproductive, but when coupled with a critical appreciation of the final product may be just what

is needed in order to solve problems in a new way. This is characteristic of genius, and even of less exalted achievers, in both science and the arts. It seems possible that here we have the link between creativity and intelligence.

We thus have a restriction of the "blind-variation-and-selection" model along two lines. The concept of g_c is defined by the number of elements available to the problem solver; the concept of creativity or originality by the *inclusiveness* of the problem solver's thinking. These limitations are probably linked with important personality traits; g_c is closely related to introversion (Eysenck, 1978; Eysenck, & Cookson, 1969; Wankowsky, 1973). This no doubt is the reason that successful scientists tend strongly to be introverted (Cattell & Butcher, 1968; Roe, 1951).

This clearly is not the place to explicate further the complexities of our theory (Eysenck, 1993b). It may, however, be useful to stress the importance of supplementing any *cognitive* theory of intelligence (g) by considering the relevance to performance of personality factors. Motivation too must be considered, although its influence on IQ testing has been found weaker than expected (Eysenck, 1944). Scientific analysis aims to sort out the major variables which deserve independent study, but in the application to real-world behavior these independent influences must of course be combined in order to assess their relevance and combination. In order to achieve a proper understanding of such independent determinants, a proper, testable theory is urgently required. That offered here is only a first beginning in the search for such a theory, but a beginning has to be made somewhere if we are not to continue to drift along a theoryless byway.

REFERENCES

- Alexander, W. P. (1935). Intelligence, concrete and abstract. *British Journal of Psychology* (Monogr. Suppl.), 19, 177.
- Barrett, P., Daum, I., & Eysenck, H. J. (1990). Sensory nerve conduction and intelligence: A methodological study. *Journal of Psychophysiology*, 4, 1–13.
- Barrett P., & Eysenck, H. (1993). Sensory nerve conduction and intelligence: A replication. *Personality and Individual Differences*, 15, 249–260.
- Barrett, P., & Eysenck, H. (in press). The relationship between evoked potential component amplitude, latency, contour length variability zero-orderings, and psychometric intelligence. *Personality and Individual Differences*.
- Barron, F. (1969). *The creative person and the creative process*. New York: Holt.
- Barron, F. (1972). *Artists in the making*. New York: Seminar Press.
- Berger, M. (1976). *A study of the Eysenck-Furueaux approach to the analysis of performance on intelligence tests*. Unpublished PhD thesis, University of London.

- Berger, M. (1982). The "scientific approach" to intelligence: An overview of its history with special reference to mental speed. In H. J. Eysenck (Ed.), *A model for intelligence*. New York: Springer-Verlag.
- Binet, A. (1903). *L'Etude expérimentale de l'intelligence*. Paris: Schleicher, Frenes.
- Binet, A. (1907). *La psychologie du raisonnement*. Paris: Alcan.
- Boring, E. G. (1957). *A history of experimental psychology*. New York: Appleton-Century-Crofts.
- Brierley, H. (1961). The speed and accuracy characteristic of neurotics. *British Journal of Psychology*, 52, 273–280.
- Buktenica, N. A. (1971). Auditory discrimination: A new assessment procedure. *Exceptional Children*, 38, 253–261.
- Burt, C. (1909). Experimental tests of general intelligence. *British Journal of Psychology*, 3, 94–177.
- Burt, C. (1940). *Factors of the mind*. London: University of London Press.
- Callaway, E. (1975). *Brain electrical potentials and individual psychological differences*. London: Grune and Stratton.
- Cameron, N. (1939). Schizophrenic thinking in a problem-solving situation. *Journal of Mental Science*, 85, 1012–1035.
- Cameron, N., & Margaret, A. (1951). *Behavior pathology*. Boston: Houghton Mifflin.
- Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychology Review*, 67, 380–400.
- Carlson, J. S., & Widaman, K. F. (1987). Elementary cognitive correlates of G: Progress and prospects. In P. A. Vernon (Ed.), *Speed of information processing and intelligence* (pp. 69–99). Norwood, NJ: Ablex.
- Carroll, J. B., (1993). *Human cognitive abilities*. Cambridge, UK Cambridge University Press.
- Carroll, J. B., Kohlberg, L., & De Vries, R. (1984). Psychometric and Piagetian intelligence: Toward resolution of a controversy. *Intelligence*, 8, 67–91.
- Cattell, J. M. (1890). Mental tests and measurements. *Mind*, 15, 373–381.
- 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.
- Cattell, R. B. (1980). The heritability of fluid, g_f , and crystallized, g_c , intelligence, estimated by a least squares use of the MAVA method. *British Journal of Educational Psychology*, 50, 253–265.
- Cattell, R. B., & Butcher, H. J. (1968). *The prediction of achievement and creativity*. New York: Bobbs-Merrill.
- Crutchfield, R. S. (1962). Conformity and creative thinking. In H. E. Gruber, G. Terrell, & M. Werkheimer (Eds.), *Contemporary approaches to creative thinking*. New York: Etherton Press.
- Deary, I. J., & Caryl, P. G. (1993). Intelligence, EEG and evoked potentials. In P. A. Vernon (Ed.), *Biological approaches to the study of human intelligence*. Norwood, NJ: Ablex.

- Detterman, D. K. (1982). Does "g" exist? *Intelligence*, 6, 99–108.
- 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.
- Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. *Intelligence*, 13, 349–359.
- Eysenck, H. J. (1939). Primary mental abilities. *British Journal of Educational Psychology*, 9, 230–275.
- Eysenck, H. J. (1944). The effects of incentives on neurotics, and the variability of neurotics as compared with normals. *British Journal of Medical Psychology*, 20, 207–227.
- Eysenck, H. J. (1953). *Use and abuses of psychology*. London: Penguin.
- Eysenck, H. J. (1967). Intelligence assessment: A theoretical and experimental approach. *British Journal of Educational Psychology*, 37, 81–98.
- Eysenck, H. J. (1969). Classification in science and principles of numerical taxonomy. In H. J. Eysenck & S. B. G. Eysenck (Eds.), *Personality structure and measurement* (pp. 3–10). London: Routledge and Kegan Paul.
- Eysenck, H. J. (1973a). *The inequality of man*. London: Temple Smith.
- Eysenck, H. J. (1973b). Personality, learning, and "anxiety." In H. J. Eysenck (Ed.), *Handbook of abnormal psychology* (pp. 390–418). London: Pitman.
- Eysenck, H. J. (1978). The development of personality and its relation to learning. In S. Murray (Ed.), *Melbourne studies in education* (pp. 134–187). Melbourne, Australia: University Press.
- Eysenck, H. J. (Ed.). (1979). *The structure and measurement of intelligence*. New York: Springer-Verlag.
- Eysenck, H. J. (Ed.). (1982). *A model for intelligence*. New York: Springer-Verlag.
- Eysenck, H. J. (1983). Revolution dans la theorie et al mesure de l'intelligence. *Revue Canadienne de Psycho-Education*, 12, 3–17.
- Eysenck, H. J. (1986a). The theory of intelligence and the psychophysiology of cognition. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 3). Hillsdale, NJ: Erlbaum.
- Eysenck, H. J. (1986b). Intelligence: The new look. *Psychol. Beitr.*, 28, 332–365.
- Eysenck, H. J. (1986c). Inspection time and intelligence: A historical introduction. *Personality and Individual Differences*, 7, 603–607.
- Eysenck, H. J. (1987a). Thomson's "bonds" or Spearman's "energy": Sixty years on. *Mankind Quarterly*, 27, 259–274.
- Eysenck, H. J. (1987b). Speed of information processing, reaction time, and the theory of intelligence. In P. A. Vernon (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* (Vol. II). Norwood, NJ: Ablex.
- Eysenck, H. J. (1993a). The biological basis of intelligence. In P. A. Vernon (Ed.), *Biological approaches to the study of human intelligence*. Norwood, NJ: Ablex.

- Eysenck, H. J. (1993b). Personality and creativity: Suggestions for a theory. *Psychological Inquiry*, 4, 147–178, 238–246.
- Eysenck, H. J., & Barrett, P. (1985). Psychophysiology and the measurement of intelligence. In C. R. Reynolds & V. Willson (Eds.), *Methodological and statistical advances in the study of individual differences*. New York: Plenum.
- Eysenck, H. J., & Cookson, D. (1969). Personality in primary school children: I. Ability and achievement. *British Journal of Educational Psychology*, 39, 109–122.
- Eysenck, H. J., & Eysenck, S. B. G. (1975). *The Eysenck personality questionnaire*. London: Hodder & Stoughton.
- Eysenck, H. J., & Eysenck, S. B. G. (1976). *Psychoticism as a dimension of personality*. London: Hodder & Stoughton.
- Fagan, J. F. (1984). The relationship of novelty preferences during infancy to later intelligence and later recognition memory. *Intelligence*, 8, 339–346.
- Fagan, J. F., & McGrath, S. K. (1981). Infant recognition memory and later intelligence. *Intelligence*, 5, 121–130.
- Frearson, W., & Eysenck, H. J. (1986). Intelligence, reaction time and a new “odd-man-out” R. T. paradigm. *Personality and Individual Differences*, 7, 807–817.
- Frearson, W., Eysenck, H. J., & Barrett, P. T. (1990). The Furneaux model of human problem solving: Its relationship to reaction time and intelligence. *Personality and Individual Differences*, 11, 235–257.
- Furneaux, D. (1952). Some speed, error, and difficulty relationships within a problem solving situation. *Nature*, 170, 37.
- Furneaux, D. (1961). Intellectual abilities and problem solving behaviour. In H. J. Eysenck (Ed.), *Handbook of abnormal psychology* (pp. 167–192). New York: Basic Books.
- Galton, F. (1883). *Inquiries into human faculty*. London: Macmillan.
- Galton, F. (1892). *Hereditary genius*. London: Macmillan.
- Gardner, H. (1984). *Frames of mind*. London: Heinemann.
- Getzels, J. W., & Jackson, P. W. (1962). *Creativity and intelligence*. New York: Wiley.
- Götz, K. O., & Götz, K. (1979a). Personality characteristics of professional artists. *Perception and Motor Skills*, 49, 327–334.
- Götz, K. O., & Götz, K. (1979b). Personality characteristics of professional artists. *Perceptual and Motor Skills*, 49, 919–924.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Guilford, J. P. (1981a). Some changes in the structure of intellect model. *Education and Psychological Measurements*, 48, 1–4.
- Guilford, J. P. (1981b). Higher-order structure-of-intellect model activities. *Multivar. Behav. Res.*, 16, 411–435.
- Guilford, J. P., & Hoepfner, R. (1971). *The analysis of intelligence*. New York: McGraw-Hill.
- Gustafsson, J. -E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, 8, 179–203.
- Guttman, L. (1954). A new approach to factor analysis: The rodex. In P. F. Lazanfeld (Ed.), *Mathematical thinking in the social sciences*. Glencoe, CA: Free Press.
- Hammer, M., & Zubin, J. (1968). Evolution, culture and psychopathology. *Journal of General Psychology*, 78, 154–175.

- 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). New York: Springer-Verlag.
- Hendrickson, D. E. (1982). The biological basis of intelligence: Part II. Measurement. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 197–228). New York: Springer-Verlag.
- Heston, L. L. (1966). Psychiatric disorders in foster home reared children of schizophrenic mothers. *British Journal of Psychiatry*, 112, 819–825.
- Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychology Review*, 75, 242–259.
- Horn, J. L. (1977). Personality and ability theory. In R. B. Cattell & R. M. Dreger (Eds.), *Handbook of modern personality theory*. London: Hemisphere.
- Horn, J. L. (1985). Remodeling old models of intelligence. In B. B. Walman (Ed.), *Handbook of intelligence* (pp. 267–300).
- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology*, 57, 253–270.
- Horn, J. L., & Cattell, R. B. (1982). Whimsy and misunderstandings of G_f ;ms G_c theory: A comment on Guilford. *Psychology Bulletin*, 91, 623–633.
- Horn, J. L., & Knapp, J. R. (1973). On the subjective character of the empirical base of Guilford's structure-of-intellect model. *Psychology Bulletin*, 80, 33–43.
- Hoyle, F. (1962). *Astronomy*. London: Macdonald.
- Hudson, L. (1966). *Contrary imaginations*. London: Methuen.
- Hudson, L. (1968). *Frames of mind: Ability, perception and self-perception in the arts and sciences*. London: Methuen.
- Humphreys, L. G., & Parsons, C. K. (1979). Piagetian tasks measure intelligence and intelligence tests across cognitive development: A re-analysis. *Intelligence*, 3, 369–382.
- Jäger, A. O. (1967). *Dimensionen der Intelligenz*. Göttingen: Hogrefe.
- Jäger, A. O. (1984). Intelligenzstrukturforschung: Konkurrierende Modelle, neue Entwicklungen, Perspektiven. *Psychol. Rundschau*, 35, 21–35.
- Jäger, A. O., & Hoenmann, H. -J. (1983). Demonstrationen von g (der allgemeinen Intelligenz) und zur Bedeutung des Variablen Kontextes bei exploratorischen Faktoranalysen. *Psychol. Beitr.*, 23, 408–420.
- Jarvik, L. F., & Chadwick, S. B. (1973). Schizophrenia and survival. In M. Hammer, K. Salzinger, & S. Sutton (Eds.), *Psychopathology*. New York: Wiley.
- Jaschek, L., & Jaschek, M. (1990). *The classification of stars*. Cambridge: University Press.
- Jensen, A. R. (1982a). Reaction time and psychometric g. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 93–132). New York: Springer-Verlag.
- Jensen, A. R. (1982b). The chronometry of intelligence. In R. J. Sternberg (Ed.), *Advances in research in intelligence* (Vol. 1, pp. 242–267). Hillsdale, NJ: Erlbaum.
- Jensen, A. R. (1984). Test validity: g versus the specificity doctrine. *J. Social Biological Structures*, 7, 93–118.

- Jensen, A. R. (1985). Methodological and statistical techniques for the chronometric study of mental abilities. In C. R. Reynolds & V. L. Wilbram (Eds.), *Methodological and statistical advances in the study of individual differences* (pp. 51–116). New York: Plenum Press.
- Jensen, A. R. (1986a). Individual differences in the high *g* paradigm. In P. A. Vernon (Ed.), *Speed of information processing and intelligence*. Norwood, NJ: Ablex.
- Jensen, A. R. (1986b). The *g* beyond factor analysis. In J. C. Connoley, J. A. Glover, & R. R. Ronning (Eds.), *The influence of cognitive psychology on testing and measurement* (pp. 87–142). New York: Plenum Press.
- Jensen, A., & Weng, L. -J. (in press). What is a good *g*? *Intelligence*.
- Karlsson, J. L. (1968). Genealogic studies of schizophrenia. In D. Rosenthal & S. S. Kety (Eds.), *The transmission of schizophrenia*. Oxford, UK: Pergamon.
- Karlsson, J. L. (1970). Genetic association of giftedness and creativity with schizophrenia. *Heredity*, 66, 177–182.
- Kaufman, A. S. (1979). *Cerebral specialization and intelligence testing*. New York: Wiley
- Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.
- Kuhn, T. S. (1974). Second thoughts on paradigms. In F. Suppe (Ed.), *The structure of scientific theories* (pp. 459–482). London: University of Illinois Press.
- Lewis, C. T., & Short, C. (1945). *A Latin dictionary*. Oxford, UK: Clarendon Press.
- Little, W., Fowler, H. V., & Coulson, J. (1947). *The short Oxford English dictionary*. Oxford, UK: Clarendon Press.
- Loevinger, J. (1951). Intelligence. In H. Helson (Ed.), *Theoretical foundations of psychology*. New York: Van Nostrand.
- Lytton, H. (1971). *Creativity and education*. London: Routledge & Kegan Paul.
- MacKinnon, D. W. (1962). The nature and nurture of creative talent. *American Psychologist*, 17, 484–495.
- MacKinnon, D. W., et al. (1961). *Proceedings of the conference on "The Creative Person."* University of California Alumni Center, Lake Tahoe, CA.
- Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence*, 7, 107–128.
- Maxwell, A. E. (1972). Factor analysis: Thomson's sampling theory recalled. *British Journal of Mathematical & Statistical Psychology*, 25, 1–21.
- McNeil, T. F. (1971). Prebirth and postbirth influence on the relationship between creative ability and recorded mental illness. *Journal of Personality*, 39, 391–406.
- Nagoshi, C. T., & Johnson, R. C. (1986). The ubiquity of *g*. *Personality and Individual Differences*, 7, 201–207.
- Neisser, V. (1979). The concept of intelligence. *Intelligence*, 3, 217–227.
- Nettelbeck, T. (1986). Inspection time and intelligence. In P. A. Vernon (Ed.), *Speed of information processing and intelligence* (pp. 295–346). Norwood, NJ: Ablex.
- Payne, R. W., & Hewlett, J. H. G. (1960). Thought disorders in psychotic patients. In H. J. Eysenck (Ed.), *Experiments in personality* (Vol. 2, pp. 13–104). London: Routledge and Kegan Paul.
- Plomin, R., De Fries, J. C., & McClearn, G. E. (1990). *Behavioural genetics*. New York: Freeman.

- Popper, K. (1935). *Logik der Forschung*. Wien, Germany: J. Springer.
- Raz, N., & Willerman, L., & Yama, M. (1987). On sense and senses: Intelligence and auditory information processing. *Personality and Individual Differences*, 8, 201–210.
- Revensdorff, D. (1978). Vom unsinnigen Aufwand. *Archiv. für Psychol.*, 130, 1–36.
- Roe, A. (1951). A psychological study of eminent physical scientists. *Genetic Psychology (Monogr. 43)*, 121–239.
- Schmidt, F. L., & Crano, W. D. (1974). A test of the theory of fluid and crystallized intelligence in middle- and low-socioeconomic status children. *Journal of Educational Psychology*, 66, 255–261.
- Shields, J. (1962). *Monozygotic twins*. London: Oxford University Press.
- Simonton, D. K. (1988). *Scientific genius*. New York: Cambridge University Press.
- Snow, R. E. (1980). Aptitude processes. In R. E. Snow, P. -A. Federico, & W. E. Montague (Eds.), *Aptitude, learning, and instruction* (pp. 27–63). Hillsdale, NJ: Erlbaum.
- Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 47–103). London: Erlbaum.
- Snow, R. E., Lohman, D. F., Marshalek, B., Yalow, E., & Webb, N. (1977). *Correlation analyses of reference aptitude constructs* (Tech. Rep. No. 5). Stanford, CA: Stanford University Aptitude Research Project.
- Snyderman, M., & Rothman, S. (1987). Survey of expert opinion on intelligence and aptitude testing. *American Psychologist*, 42, 137–144.
- Snyderman, M., & Rothman, S. (1988). *The I. Q. controversy*. New Brunswick, NJ: Transaction Books.
- Sokal, R., & Sneath, P. (1963). *Principles of numerical taxonomy*. London: W. H. Freeman.
- Spearman, C. (1904). “General intelligence”: Objectively determined and measured. *American Journal of Psychology*, 15, 201–293.
- Spearman, C. (1923). *The nature of “intelligence” and the principles of cognition*. London: Macmillan.
- Spearman, C. (1927). *The abilities of man*. London: Macmillan.
- Spearman, C., & Jones, L. (1950). *Human ability*. London: Macmillan.
- Stankov, L., & Horn, J. L. (1980). Human abilities revealed through auditory tests. *Journal of Educational Psychology*, 75, 21–44.
- Sternberg, R. J. (1982). *Handbook of human intelligence*. London: Cambridge University Press.
- Sternberg, R. J. (1985). *Beyond I. Q.* Cambridge, UK: Cambridge University Press.
- Sternberg, R. J., Conway, B. E., Ketron, J., & Bernstein, M. (1981). People’s conception of intelligence. *Journal of Personality & Social Psychology*, 41, 37–55.
- Sternberg, R. J., & Detterman, D. K. (Eds.). (1986). *What is intelligence?* Norwood, NJ: Ablex.
- Sternberg, R. J., & Wagner, R. K. (Eds.). (1986). *Practical intelligence*. Cambridge, UK: Cambridge University Press.
- Suppe, F. (Ed.). (1974). *The structure of scientific theories*. Chicago, IL: University of Illinois Press.

- Thomson, G. H. (1939). *The factorial analysis of human ability*. London: University of London Press.
- Thorndike, E. L., Bregman, E. O., Cobb, M. V., & Woodyard, E. I. (1926). *The measurement of intelligence*. New York: Teachers College, Columbia University.
- Thurstone, L. L., & Thurstone, T. R. (1941). *Factorial studies of intelligence* (Psychometric Monographs, 2).
- Torrance, E. P. (1962). *Guiding creative Talent*. Englewood Cliffs, NJ: Prentice-Hall.
- Tuddenham, R. D. (1962). The nature and measurement of intelligence. In L. Portman (Ed.), *Psychology in the making* (pp. 469–525).
- Undheim, J. O., & Gustafsson, J. -E. (1987). The hierarchical organization of cognitive abilities: Restoring general intelligence through the use of linear structural relations (LISREL). *Multivariate Behavior Research*, 22, 149–171.
- Undheim, J. O., & Horn, J. L. (1977). Critical evaluation of Guilford's structure-of-intellect theory. *Intelligence*, 1, 65–81.
- Vernon, P. A. (Ed.). (1993). *Biological approaches to the study of human intelligence*. Norwood, NJ: Ablex.
- Vernon, P. A., & Mori, M. (1989, June). Nerve conduction velocity: A physiological correlate of intelligence and speed of information-processing. *Proceedings of the 4th Meeting of the International Society for the Study of Individual Differences*.
- Vernon, P. E. (1950). *The structure of human abilities*. London: Methuen.
- Vernon, P. E. (1979). *Intelligence, heredity and environment*. San Francisco: Freeman.
- Wallach, M. A., & Kogan, N. (1965). *Modes of thinking in young children*. New York: Holt, Rinehart & Winston.
- Wankowsky, J. A. (1973). *Temperament, motivation and academic achievement*. Birmingham, UK: University of Birmingham Press Educational Survey & Counselling Unit.
- White, P. O. (1973). A mathematical model for individual differences in problem solving. In A. Elithorn & D. Jones (Eds.), *Artificial and human thinking* (pp. 246–260). Amsterdam: Elsevier.
- White, P. O. (1982). Some major components in general intelligence. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 44–90). New York: Springer-Verlag.
- Wissler, C. (1901). The correlation of mental and physical tests. *Psychological Review* (Monogr. No. 3).
- Wolman, B. B. (Ed.). (1985). *Handbook of intelligence*. New York: Wiley.
- Woody, E., & Claridge, G. (1977). Psychoticism and thinking. *British Journal of Social & Clinical Psychology*, 16, 241–248.
- Wright, G. H. (1971). *Explanation and understanding*. Ithaca, NY: Cornell University Press.

Chapter 7

*A System for Understanding
Cognitive Capabilities: A
Theory and the Evidence on
Which it is Based*

John Horn and Jennie Noll

University of Southern California

PURPOSES

We present a theory of human cognitive capabilities—a theory of what is often (as in this book) referred to as intelligence. We use the term *cognitive capabilities* rather than *intelligence* to refer to these phenomena in an effort to avoid begging an important question about the nature of the phenomena. The singular word *intelligence* implants the assumption that the phenomena are unitary—that a single concept represents the entirety. But the question of whether or not there is such unity is a major question, the answer to which should be adduced from carefully accumulated evidence. This is not a matter of assumption: it is a matter of what is the nature of the phenomena that are referred to with the word *intelligence*. These phenomena are human capabilities referred to with such terms as thinking, problem solving, memory, abstracting, reasoning,

retrieval, concept attainment, knowing, and behaving intelligently. The question is: "Are these phenomena well explained with one concept, intelligence; are the phenomena unitary?" The answer appears to be "No."

One may ask: "Well, okay, but what do you mean by unitary?" We mean that the elements of that which is unitary change together systematically; they rise together, they fall together, they appear together, they disappear together. Also, there is lawful system in this change. For example, the change in one element may systematically follow change in another element, or occur only when other elements change in a particular order. There are many ways in which functional unities of the human body are systematic.

There are a number of ways in which evidence might be accrued to show that different abilities constitute a unity—intelligence. Evidence might show that all of a set of abilities are caused by a unitary set of genetic determiners. The term *intelligence* is often used in a manner that suggests that such evidence is at hand.

In fact, however, we still have much to learn about genetic determiners of abilities, and the evidence thus far accumulated does not indicate a unitary trait. The evidence points to several sets of genes separately and only partially determining different cognitive abilities. No particular configuration of these abilities constitutes intelligence. It is rather as if there are several intelligences, not one. Different cognitive abilities are analogous to different features of physiognomy. Just as different sets of genes determine the shapes of ears, the breadth of noses, the color of hair, and other such features that make up the different looks of different people, so different sets of genes determine the physiological structures and functions that underlie the developments of different cognitive capabilities that make up the different intelligences of different people. The different intelligences do not appear to be inherited as a unitary whole.

Similarly, the evidence indicating determinants in learning of abilities points to many factors of motivation, timing, practice, reinforcement, and developmental stage that partially determine the acquisition of the abilities that make up what is referred to as intelligence. These determinants do not work in unison to produce a single unitary whole. The different organizations of distinct cognitive capabilities that emerge from learning indicate different intelligences. There appears to be no single intelligence that is an outcome of learning.

The different abilities of any one person are contained within and are part of the functioning of that person alone and thus are unitary in this very broad sense. But in this sense all psychological phenomena are unitary—motivation, temperament, belief, and so on. In scientific description and explanation we identify distinct functional unities within the unities of separate individuals. The heart is a unity separate from the liver and the brain, for example. There are notable individual differences in hearts and livers and brains (and parts of brains). There are notable individual differences in the development of the liver

and the heart. The features of the heart change together in an interrelated manner that is distinct from changes in features of the liver. So it is that separate abilities may be distinct unities. A fundamental scientific question is: Are different cognitive capabilities indicative of distinct structures and developments, or, alternatively, is there a single system of relationships among all these structures and developments such that one can justify use of a single unifying concept—intelligence? We should not assume the answer to this question: We should gather evidence and study it carefully to arrive at the answer. The purpose of this chapter is to present the evidence (such as we can comprehend it) that is relevant to this answer.

The theory we will describe in this chapter is in some respects different from other theories dealing with the same phenomena, but we seek to avoid controversy that is based on opinion or equally defensible interpretations of the same evidence. The theory is based on the assumptions of the paradigms of the research from which the evidence has been obtained (in the broad sense that Kuhn, 1970, defines paradigm). Any science is limited by such assumptions. We will endeavor to make these assumptions evident, but not by detailed philosophical analysis; rather by clear description of the nature of the evidence.

Given these assumptions, we aim to present only evidence and reasonable (plausible) interpretation of the evidence. Our major purpose is to provide a conceptual basis for further scientific research on the nature and development of human cognitive capabilities.

PROBLEMS

When one looks for a studied moment at the myriad of abilities that humans display, it's as if one were to look into the heavens on a clear night and become stirred by the ceaseless drift of the clouds of the Milky Way. On such a night one might be dimly aware that there is order and system in the celestial white. But where among the drifting haze might one draw dimensions to represent this order? At first there is no answer to this question, only befuddlement. The same is true for human abilities. They appear as free-floating swarms emerging from spaces of unknown many dimensions. Is there genuine order in this throng, or can one at least impose an order that will not do great injustice to the complexity and still enable one to organize thinking and talking about it? (Horn, 1988, p. 645)

It is tempting to suppose that before there can be sound research on human intelligence the phenomena must be definitively circumscribed—and on this basis set out to try to specify the domain of the abilities of intelligence. The problem is that this can't be done. It cannot be done not only because the phenomena are so vast and diverse, as indicated above, but also because they are constantly changing. The swarms of human capabilities that indicate intelli-

gence are ever-expanding. New capabilities appear with every invention of the human mind—the printing press, the silicon chip, DNA, computer games, fusion and fission, and so on, ad infinitum. The set of abilities that indicate intelligence is also shrinking. Abilities that once were important are no longer developed, perhaps because the culture no longer needs them. Efforts to circumscribe the domain of intellectual capabilities thus are doomed to failure, because not only is the universe of such elements so vast that its boundaries are beyond comprehension, but also because the domain of such capabilities is constantly growing and shrinking, evolving into a new vastness.

Not only is the universe of human cognitive capabilities vast and diverse, there are also no clearly discernable demarcations to distinguish abilities one from another or even from other qualities of the human, such as those of motivation and temperament. There is no clean way to divide the manifest phenomena into discrete segments. Look at it this way: What do you see when you see ability in action? You see abilities—many of them, perhaps all of them—and you see many other things as well—emotion, drive, apprehension, feeling, and so on. You do not see these “things” as discrete elements: You see them as a seamless whole. You do not see a line in behavior between reasoning, retrieval, perception and detection. Thus it is that abilities are not segments of behavior; they are abstractions we apply to an indivisible flow of behavior. To ask “Is there order among swarms of human abilities?” is to ask what concepts, and how many, are needed to describe an indivisible whole of human capabilities. In specifying different features of cognition, it is as if we are slicing smoke. With our methods, we slice a continuous, homogenous, irregular mass of gray into shades of gray; with our words, we then describe the segments thus cut.

We cannot hope to define human capability as a concrete entity as we define a brain or a heart. But there is behavior that people—scientists and others alike—refer to as indicating cognitive capabilities. This behavior can be objectively and accurately recorded. Operational definitions of concepts representing this behavior can be specified in measures that are reliable and reliably distinct one from another. Samples of the behavior that scientists agree is indicative of cognitive capabilities can be obtained with these measures in goodly samples of humans. The regularities among these measures can be identified. Explanation of these regularities—scientific theory—thus can be educed. Thus, we can understand human cognitive capabilities in terms of operational definitions for specifying behavior and awareness of regularities—repeatable relationships—that indicate how these definitions are associated with other variables of function and development. This is the task of building construct validity (Cronbach, 1977, 1987). It is the task of establishing verified scientific theory.

NUMBER

We thus come to the question of how many human abilities are there: How many shades of gray abilities can we see? The answer is not simple.

From his studies of highly talented people, Commons (1985, p. 2) estimated that "There might be 800,000—or more intellectual abilities!" Carroll has told us, on the other hand, that "it has become clear that there exist only a relatively small number of identifiable, replicable abilities" (Carroll, 1985, p. 9). Who is correct? The answer is both.

Carroll's correct statement is a conclusion drawn from thoughtful analysis of a large body of replicated results of studies in which large samples of notably different people performed on large batteries of diverse tests and paradigms, each designed to measure a feature of human intelligence—a process, a capacity, a quality, an ability—that is different from the features measured by other tests and paradigms. Analyses of such information indicate that measures designed to indicate distinct abilities are not distinct. The measures overlap. They have many features in common: A relatively small number of common factors will account for the covariability among the performances on many different tests and paradigms.

An ability is common to several tests if the order of subjects on measures of that ability is the same (except for error, of course) as the order obtained by summing over each subject's scores on the several tests (thus obtaining an overall measure of each subject). If few such abilities will accomplish this for all possible sets of "several tests" within a large battery of tests, then it can correctly be said that a small number of factors measure abilities that are common to the large number of tests. The evidence of hundreds of studies indicates that indeed, a small number of factors will measure all the abilities that are common in very large batteries. This is the sense in which there are "only a relatively small number of identifiable, replicable abilities."

The same analyses that indicate few common factors also point to the many abilities to which Commons refers. These analyses indicate many nonerror specific factors. These are reliable (replicable) measures of the behavior evinced with each particular test—segments not correlated with (predicted by, accounted for by) the other tests and paradigms of a battery. Specific factors thus represent abilities (aspects of abilities) that are not the abilities held in common with other tests. There can be more than one such ability indicated by the specific factor of each test. Since many, many tests (to measure a very large number of human abilities) have been invented, many, many specific factors of ability have been indicated by the evidence that also indicates few common factors. Many more such tests and paradigms can be—and will be—invented. As circumstances of human existence change, new abilities, and new tests to

measure abilities, will be developed. There is virtually no limit to the new tests that humans can invent to measure new abilities (Humphreys, 1979). Thus it is that there can be many, many specific abilities. When one realizes, also, that every human possesses billions of neurons that can be arranged in billions of ways, and each arrangement can support a different pattern of abilities, then again it is not difficult to suppose that there are, as Commons observed, perhaps 800,000 human cognitive abilities.

Thus it is that there are both many and few human abilities. But it is not possible to understand all of the “many.” In seeking scientific understanding we are driven to deal with fewer than “many.” Research has been directed at describing the “few” that account for the most—namely, the common abilities to which Carroll referred.

A common factor indicates unity in the sense that the measures of the different capabilities of the factor rise and fall concomitantly in individual differences. A system of common ability factors is a system of such unities.

A common factor unity does not necessarily indicate functional unity—in the sense that all components of the factor change as a system under all conditions of function—but it provides a *prima facie* case for such unity. It is thus a starting point. Covariation in one media of observation is often indicative of covariation in other media. Covariation in individual differences can thus be indicative of covariation in change through time (e.g., of development), change with alteration of structure (e.g., brain damage), or change with treatment (e.g., of educational program). A functional unity is indicated by evidence of systematic covariability in all of such different media of observation. Abilities indicated by common factors provide a rational basis for research aimed at identifying such systematic covariability.

The system of common factor abilities that provides the basis for the modern theory of cognitive capabilities developed over a period of the first 90 years of this century. It is useful, in thinking about the theory and the assumptions on which it is based, to have some awareness of the history of development of the theory. To this we shall turn next.

ROOTS

The idea of general intelligence became prominent in the early part of this century. That idea was gradually replaced, in response to the evidence of empirical research, by the idea of several intelligences.

Spearman (1937) traced ideas about human capabilities from the age of Pericles into the present century. Through the ages, there have been attempts to describe features of thinking that distinguish humans from other animals and distinguish particularly wise and capable individual humans from others not so wise and capable. Features frequently identified include ability to abstract, to

adapt, to anticipate, to combine ideas in novel ways, to quickly comprehend, to conceive universal ideas, to educe relations, to make inferences, to form categories, to generate ideas, to form concepts, to use language, to draw valid conclusions, and so on. The listing could go on. Many, many features have been discussed. Spearman identified theories of multiple abilities, theories of a few major abilities, and theories of one major ability.

While ideas about human cognitive capabilities are of ancient origin, the measurement of abilities—with the repeatable operations that define objective tests—became a reality only in this century. There were demonstrations in the 1890s that responses to questions (test items) could be grouped together to provide indications of memory, knowledge, verbal comprehension, and reasoning. The first measure of intelligence was based on these demonstrations. Very likely you know that story—about how the French Ministry of Education asked Binet, and presumably his coworker Simon, to devise a means for distinguishing between children who were slow in learning and children who were normal learners in the early years of schooling; how Binet and Simon (1905) put together a diverse set of questions based on their studies of memory, verbal comprehension, and so on to form a single composite test; how they developed a measure from the sum of the number of correct answers a child gave (relative to the average of children that child's age); and how this measure of a mixture of abilities came to be known as intelligence. Perhaps you know the story, too, of how with the suggestion of Stern (1914) that a ratio of performance score to the child's chronological age be calculated; this was called the intelligence quotient, shortened to IQ. The Binet-Simon test proved to be useful for making the placement decisions of the French Education Ministry's request. It provided information about the abilities of children that otherwise had to come from judgments by skilled teachers. The judgmental measures were subjective and expensive; the measures obtained with the test were more nearly objective and relatively inexpensive. The test worked for the purpose for which it was designed. It became the benchmark and model for construction of almost all of what subsequently were called intelligence tests. A test developed in this country by Terman (1916) at Stanford University came to be particularly widely used; The Stanford-Binet test is still widely used today.

With the success of the Binet-Simon kind of measure came much thought about just what the test measured and what could be wrong, as well as right, about decisions based on the measurement discriminations it provided. A great amount of effort was devoted to defining (in words) the IQ this kind of test was said to measure. These definitions and various applications of the test generated much controversy. New tests, designed to be improvements, were constructed. These generated more efforts to define and more controversy. The diversity of intelligence tests and definitions of intelligence led some to opine that intelligence is nothing more than whatever it is a particular intelligence test measures (Boring, 1923).

But all was not simply a flurry to define intelligence in words, redo the old test, and describe new tests of the Binet-Simon variety. At the same time the Binet-Simon test came into use, Spearman put forth a detailed theory of what tests measured and how all cognitive tests could involve the education or relations and correlates. This theory and other such analyses and questioning of what was measured led to hypotheses; many experiments were designed to refine understanding of the capabilities that were expressed in various tests; old theories were modified and new theories were developed to explain results from these experiments. A wide diversity of scientific theories grew out of this movement (Sternberg & Detterman, 1986).

The diversity of theories of cognitive capabilities creates confusion and frustration. We will try to dissolve some of this confusion and frustration by organizing theories. First, we will organize theories of a single general intelligence into four classes: essence, compound, mixture, and one-test theories. Then we will describe multiple intelligence theories in a context of historical development.

THEORIES OF GENERAL INTELLIGENCE

As we have seen, the conglomerate measure that Binet and Simon developed was interpreted as indicating a general capability labeled IQ and interpreted as general intelligence. Spearman (1904, 1927) concluded from his historical review and from a series of empirical studies that there was one major cognitive capability, which he labeled *g*, a symbol Spearman selected in an effort to get away from unwanted connotations of the term *intelligence*. The intent of this labeling did not work: The *g*-factor was interpreted as intelligence. The Binet-Simon and Spearman studies were particularly influential. Theories of a general factor became entrenched in the thinking and research of behavioral scientists of this century. Spearman's theory is an example of an essence theory; the Binet-Simon theory represents mixture theories.

Essence Theory

Essence theories of intelligence stipulate that all distinct intellectual abilities stem from one basic process, one element—the essence of intelligence. Depending on the theory, the essence may come from any of several possible sources. It might be argued that the essence is inherited, for example, in which case the essence would be embedded in the genetic structure and function. But an essence theory could also stipulate that the essence stems from the environment. The main feature of essence theory is not from whence stems the essence, but the idea that one thing—the essence—determines what we see.

Spearman's essence theory is a marvelous example of good scientific theory. The fate of the theory exemplifies how good theory in science—testable theory—is found to be inadequate and is replaced by improved theory.

The essence of g in Spearman's theory is a capacity for comprehending relationships among perceptions—what Spearman called education of relations—and a capacity for drawing accurate conclusions from awareness of relations—what he called the education of correlates. A critical hypothesis derives from the theory. Spearman derived a clearly stated mathematical and statistical model for testing this hypothesis. The hypothesis can be unambiguously falsified with a finding that data do not fit this model.

The critical hypothesis specifies that two critical features characterize the intercorrelations among all tests that researchers can agree are measures of individual differences in human cognitive capabilities: (a) that there is one, and only one, common factor among these correlations—this indicating the education of relations and correlates—and (b) that for each individual test, there is one specific factor, unique to that test, indicating a specific ability. The correlation between any two tests results from the common factor. The unique factors of different tests are uncorrelated and thus produce none of the observed test intercorrelations. The correlations among tests needn't be large; the requirement of the model is not that there be large correlations, only that the observed correlations be due to the one common factor and only to that factor. The theory thus specifies a great many human abilities—the specific abilities—each unrelated to the other after g , an ability common to all tests, has been parialed out of the relationships among the tests.

If a sample of tests that can be seen to measure cognitive capabilities is used to measure a sample of humans, the intercorrelations among these tests are obtained, and the one-common-factor model is found to fit these intercorrelations (to within chance variation), the result is evidence in support of the central hypothesis of the theory of g . This evidence alone would not indicate complete understanding of the essence any more than the identification of black holes in the universe indicates that we understand these phenomena, but the finding is extremely important because it indicates where we must seek understanding.

Shortly after Spearman (1904) proposed his theory evidence was adduced that failed to support the central hypothesis. Burt (1909, 1911) found that numerical (N) and verbal (V) common factors (called group factors), in addition to a general factor, were necessary to account for the intercorrelations among measures of achievement that had been analyzed by Spearman. Burt's criticisms and results were repeated by several investigators between 1910 and 1920. There followed a steady increase in the number of common factors found to be necessary to describe the covariance among measures accepted as indicating human cognitive capabilities. In 1924 Burt reported substantial support for common factors of memory span (Ms), manipulative ability (km), and

scholastic ability (Sc), as well as for V, N, and a general factor. To this list was added, still in the 1920s (before Thurstone, 1935), spatial (S), perceptual speed (P), mechanical reasoning (Mk), and visualization (Vz) and several less clearly indicated common factors (Alexander, 1935; Brown, 1932; Brigham, 1932; Cox, 1928; El Koussy, 1935; Kelley, 1928; Patterson & Elliot, 1930). The results from these studies, as well as from studies very carefully designed to support Spearman's theory of g (Alexander, 1935; Cattell, 1940; El Koussy, 1935; Eysenck, 1982; Rimoldi, 1948; Willoughby, 1927), demonstrated convincingly that the model for g did not fit the data of cognitive capability measures. Efforts to modify g theory by respecifying the essence (Alexander, 1935; Cattell, 1940; El Koussy, 1935; Eysenck, 1982; Rimoldi 1948; Willoughby, 1927) were unsuccessful. No support for a g theory could be found for all the many abilities that were said to indicate intelligence. More than one common factor was needed to account for the intercorrelations among different measures of such capability.

In recent years there has been a resurgence of essence theory, particularly in the work of Eysenck (1982) and Jensen (1982), but as in earlier work, the evidence does not support these new versions of the theory. The central element specified by Eysenck and Jensen is a neural mechanism—speed of neural processing—represented in behavioral operational definitions by measures of capacity for holding separate ideas and relations in the span of immediate awareness. There is little doubt that such memory is an important cognitive capability but the evidence of many studies now makes it clear that such a capacity is not the essence of all other cognitive capabilities that can be seen to be indicative of human intelligence.

The evidence accumulated throughout this century thus indicates the inadequacy of any essence theory of general intelligence. Compound theories have suffered the same fate, but they present a potential for salvaging important features of Spearman's theory.

Compound Theories of g

Spearman's substantive theory and mathematical model are very demanding. An investigator must understand the concept of eduction of relations and correlates very well to be able to put together tests that will measure only that one capacity and an ability specific to each test (and that test alone). It is easy to introduce what Spearman called *swollen specifics*—the same specific factor measured in more than one test—into a sample of tests designed to indicate g . When the same factor is measured in more than one test it becomes a common factor. For example, if g is indicated in capacity for holding elements in awareness (short-term memory) and in capacity for attaining concepts, but a battery of tests designed to indicate g has three measures of short-term memory and

five tests of concept attainment, a *g*-factor model will not fit the data because the data will indicate that there is a common factor of short-term memory and a common factor of concept attainment. On the other hand, if a battery were very carefully put together to involve only one test of short-term memory and one test of concept attainment, the *g* model could be shown to fit the data. The two abilities in this case—short-term memory and concept attainment—could be said to indicate a compound, much in the way that hydrogen and oxygen indicate a compound of water.

In chemistry a compound is a particular union of elements. The elements are different, but they combine in a specified manner to form a unit. In every quantity of water, there is a union of the elements of hydrogen and oxygen, each in a precise proportion of amount—two moles of hydrogen to one of oxygen. The necessary and sufficient conditions for a definition of water are specified in H_2O . Until a compound is broken apart, and this requires extraordinary actions, it functions as a unit. Water behaves as an entity. It is not easily divided into hydrogen and oxygen. It heats up and cools down as a whole, not as hydrogen and oxygen separately.

A theory that intelligence is a compound of separate processes can be tested with the same kind of one-common-factor model that Spearman specified for his essence theory. The difference is that each measure of a compound theory is designed to measure a different element of the compound, whereas all measures of an essence theory are designed to measure the same thing. In a compound theory it is not that all tests indicate eduction of relations and correlates; tests indicate the several elements that constitute the compound.

To demonstrate *g* as a compound, one might put together the following battery of tests:

- sensory detection, measuring immediate apprehension
- simple span memory, measuring immediate retrieval
- backward span memory (or complex reaction), measuring working (holding) memory (or span of immediate awareness)
- uses for objects, measuring retrieval from long-term storage
- common word analogies, measuring associative reasoning
- letter series, measuring inductive reasoning (eduction of relations)
- matrices, measuring deductive reasoning (eduction of conjunctive correlates)
- disjunctive concept formation, measuring eduction of disjunctive correlates

Each of the processes measured by these tests has been described as indicating intelligence or an important feature of intelligence (Carroll, 1993; Detterman, 1989; Eysenck, 1982; Horn, 1980a; Jensen, 1982; Sperling, 1960; Sternberg, 1985; Sternberg & Detterman, 1986; Woodcock, 1990). But each test has been carefully selected to not measure anything that is measured by

any of the other tests: Each measures (by hypothesis) a separate element of a compound. The sensory detection measure (Broadbent, 1966; Sperling, 1960) measures only that awareness one attains for only a split second, as in tachistoscopic presentations (Biederman, 1992), not the ability to retrieve information, as in simple span memory. A good test of working memory, on the other hand, measures only that short span of information one can hold in immediate awareness as one does other things, such as reasoning, not the clang memory that is not retained for work, as indicated by recency recall in serial learning (Glanzer & Cunitz, 1966). Each test thus is designed to measure a specific factor that is not measured by any other test.

A common factor among such measures indicates the necessary working together of separate processes. One must apprehend before one retains; one must retain before one can hold information in awareness; one must hold information in awareness in order to perceive relationships among elements of information; one must apprehend such relationships before one can inductively comprehend relations; one must comprehend relations before one can draw conclusions (as in drawing conjunctive or disjunctive deductions). It is necessary if this common factor (in linear factor analysis) is to be indicated among measures of individual differences that there be monotonic (approximately linear) relationships among the processes: Individuals high in one process, relative to individuals scoring low, must be high in the other processes.

If these conditions obtain—each test does measure a separate process essential to *g*, and whatever else a test measures is not measured in any other test of the battery—then a one, and only one, common-factor model will fit the data and the evidence of this finding supports an essence theory of intelligence that involves the indicated processes.

A one-factor model can be shown to come close to fitting data that are very carefully selected in the manner indicated (Horn, 1972b, 1980b, 1981; Horn, Donaldson, & Engstrom, 1981). Residuals left after extraction of one factor are very small. The model does not quite fit, however; the residuals are not evenly distributed around zero and a second small factor must be extracted to adequately account for the correlations. But the evidence suggests that support can be obtained for a limited one-factor theory of intelligence. Evidence from hierarchical analyses (Gustafsson, 1984, 1985; Undheim & Gustafsson, 1987), to be reviewed shortly, also provides support for such a limited theory.

Such a theory truly is very limited, however. Many abilities that have been described as indicating intelligence and important features of intelligence are not represented among those selected to indicate one common factor. Missing from this set are abilities of verbal comprehension, insight, abstraction, Gestalt closure, breadth of information, visualization, arithmetical skill, quantitative reasoning, and many others. When some of these measures are entered in the sample, a one-common-factor model does not come at all close to fitting the data. Among a goodly sample of such measures, at approximately the level of

breadth of the one factor described above, there are nine common factors. One of these is similar to the one factor described above. This is the factor that has become known as fluid intelligence, or simply Gf (Horn & Cattell, 1966). In a goodly sample of putative measures of intelligence, Gf accounts for only a small proportion of the common variance, and most of the tests described above as indicating one common factor are also significantly related to at least one common factor in addition to Gf. In the Gustafsson (1985) and Undheim–Gustafsson (1987) analyses, Gf is perfectly correlated with a single general factor in a hierarchical model, but this factor also accounts for only part of the common variance among all of a set of measures selected to measure important aspects of intelligence.

Thus, support for a hypothesis of one common factor can be adduced, and a case can be made that this factor represents a substantive theory of intelligence. But the theory is limited because it represents only a few highly selected indicators—not fully representative—of all the capabilities that scientists and others alike have determined are indicative of human intelligence.

Support for one common factor can be adduced for other carefully chosen selections of tests; that is, other than those described above for a Gf compound theory. And the selection can be such that people agree that the tests are indicative of intelligence. In this way, one can make a case for a theory of intelligence that is different from the one indicated above. But this theory, too, would be limited in the same manner as the theory of Gf is limited. The evidence for such a theory can be useful, but it does not support a theory of general intelligence. Indeed, as indicated before, diligent efforts to find support for this theory with model fitting analyses have repeatedly turned up evidence to the contrary: Indicators of intelligence simply are not interrelated in a manner that would indicate one common factor.

Mixture Theories

In mixture theories of intelligence, a mix of different abilities provides an operational definition. The mixed abilities usually are positively intercorrelated, and positive intercorrelations may be required of the mixture, but no model or rejectable hypotheses are specified for the relationships among the abilities of the mixture.

Many mixture measures of intelligence have been proposed. Most tests that are said to measure general intelligence or IQ are mixture measures. Well-known examples are the Wechsler tests and the Stanford-Binet.

Mixture measures have no intrinsic features that indicate that one is more nearly correct than another. They involve different abilities, mixed in different proportions. Depending on which abilities they include and do not include they are better and worse for different applications. They can be judged in terms of

technical criteria, such as reliability, and in terms of relevance for the decisions for which test measures might be used. The mixture of abilities best designed to indicate success in graduate study in chemistry, for example, is notably different from the mixture best designed to facilitate guidance or selection of students entering music education.

In almost any sample of people older than 5 years of age, almost all tests that reliably measure a cognitive ability correlate positively with all other such tests. The correlations may not be large, but they are almost always positive. This is referred to as the principle of positive manifold. Exceptions to this principle are rare. Almost always they are for measures obtained with highly speeded tests of trivial difficulty.

It is sometimes reasoned that positive manifold supports a theory of general intelligence. Jensen (1982, 1984) is one of the leading exponents of this position. It is reasoned that positive manifold indicates that one thing must be measured by all tests, and that almost any mixture of such tests will provide measure of this one thing. The sum of scores on any broad collection of cognitive tests, Jensen argues, will provide a good "working definition" of Spearman's g .

A major problem with this reasoning is that different mixtures measure different collections of abilities, and the different collections have different relationships with other variables. Even if all ability tests measured one thing (say, g)—and the evidence of positive manifold does not necessitate this—they measure other abilities as well, and these other abilities are in different proportions of the total in different mixtures. The orders of people from high to low on different mixture measures are different. Decisions made on such orders thus will vary with the mixture used. It is rather like thinking of mixtures of orange juice and vodka, milk and honey, and benzine and gasoline as measures of a q factor of liquidity. Such mixtures would measure one thing at an abstract level, but the construct validities for the mixture measures can be expected to be quite different.

Humphreys (1979, p. 106) defined intelligence as a mixture of "the entire repertoire of acquired skills; knowledge, learning sets, and generalization tendencies considered intellectual in nature that are available at any one period of time." He proposed that a measure of intelligence should be a representative sample from this repertoire. This is thus a sampling theory. To the extent that different mixture measures are based on representative samples of the abilities of intelligence, they measure the same intelligence.

Thomson (1919, 1948) proposed a sampling theory as an alternative to the Spearman (essence) theory of g . He reasoned that intellectual tests measure somewhat different samples of elementary capacities (referred to as bonds). The universe of such capacities is intelligence. A common factor among different tests would indicate samples of bonds representative of the universe of such capacities.

The major problem with sampling theories of intelligence is one of circumscribing the universe of abilities that make up intelligence. How does one draw, and how can one know when one has drawn, a representative sample of the entire repertoire of abilities to which Humphreys referred? If we can neither designate the population of abilities for such sampling nor specify the criteria with which to evaluate the extent to which a sample is representative of such a population, the samples of abilities that are drawn for mixture measures of general intelligence are arbitrary.

A second problem with sampling theories of intelligence is that the number of abilities that must be sampled to form a mixture that is representative (in the sense of Humphreys's or Thomson's definition) is probably large—too large to hope to sample. Thousands of human abilities have been identified; thousands more can be identified. There is no clear way to representatively sample from this large universe of abilities.

Different mixture measure tests of intelligence measure different things. Indeed, mixtures within the same test used at different ages measure different things. The Stanford-Binet for infants, for example, measures abilities that are different from those measured with the Stanford-Binet for 13-year-old children. No mixture measure test is known to be representative of the entire repertoire of acquired skills, knowledge, learning sets, and generalization tendencies considered intellectual in nature.

One-Test Theories

Rather than deal with the evidence that different abilities do not indicate the same intelligence, or use mixture measures, a few investigators have simply made up, or used, a single test to represent their particular ideas about intelligence. The Porteus (1946) maze and Raven (1938) matrices tests are examples of this response to the evidence. Such tests are very narrow samples of the phenomena investigators identify as indicating human intelligence.

STRUCTURAL THEORIES

The discovery that intercorrelations among abilities are almost always positive was made early in this century and repeatedly confirmed. But a first factor representing this positive manifold would not account for the common variability. As the evidence accumulated to indicate that a one-common-factor theory of intelligence would not suffice, multiple-factor theories were developed. Three kinds of such theories have been particularly influential: (a) hierarchical, (b) primary mental ability (PMA), and (c) the theory of fluid (Gf) and crystallized (Gc) intelligences, now in expanded form better known as Gf–Gc theory.

The models for each of these kinds of theories account for covariability additional to that accounted for with one common factor. But some models specify that one factor should account for as much of the positive manifold common variance as possible. These models are said to be *hierarchical*. On the other hand, the models for PMA theory, Gf–Gc theory, and expanded Gf–Gc theory do not require one factor to account for as much common variance as possible; usually, all of the abilities of these theories—the first as well as the others—are thought to be about equally representative of the common variance.

Hierarchical Theories

Hierarchical theories are based on two forms of multiple-factor analyses. In the first form of analysis, a general factor is extracted first and retained; in the second form, a general factor is calculated on the basis of covariability among factors such as those of PMA or Gf–Gc theory.

In the early studies following Spearman, the broadest possible factor was extracted first and usually interpreted as indicating general intelligence. Additional factors were calculated without appreciably altering this first factor. Usually, a necessary bipolarity among these residual factors was retained, so that abilities were seen as contrasts. For example, residual factor might be found to be correlated positively with verbal abilities and negatively with mathematical abilities, so the factor was seen as indicating opposition between verbal and quantitative abilities—after general intelligence was controlled. If one spent intellectual resources developing verbal abilities, the theory suggested, this took some resources away from developing quantitative abilities.

The idea that abilities are in opposition generally did not meet with favor, largely because the abilities found to be in opposition on a bipolar factor in one study were not found to be thus opposed in other studies: The results of bipolar factor analysis did not replicate dependably. Other methods were developed in which the contrast of bipolar residual factors was eliminated by reducing the variance of the general factor and calculating two factors in place of each bipolar factor (Holzinger & Harman, 1941). Most of the hierarchical theories developed in the first two-thirds of this century were based on models of this kind (Burt, 1949; Vernon, 1961).

Burt's (1941, 1949) theory of subdivided hierarchically organized abilities was particularly influential in guiding research. In this theory the mind was seen to be

organized on what can be called a hierarchial basis [in which] the processes of the lowest level are assumed to consist of simple sensations or simple move-

ments, such as can be artificially isolated and measured by tests of sensory 'thresholds' and by the timing of 'simple reactions.' The next level includes the more complex processes of perception and coordinated movement, as in experiments on the apprehension of form and pattern or on 'compound reactions.' The third is the associative level—the level of memory and of habit formation. The fourth and highest of all involves the apprehension or application of relations. 'Intelligence,' as the 'integrative capacity of the mind' is manifested at every level, but these manifestations differ not only in degree, but also (as introspection suggests) in their qualitative nature. (Burt, 1949a, p. 46)

Much of this thinking has been incorporated in Gf–Gc theory and its derivatives.

In recently developed hierarchical theories (Carroll, 1989; Gustafsson, 1985; Hakstian & Cattell, 1978; Horn, 1965, 1972a; Undheim & Gustafsson, 1987) the general factor is specified on the basis of intercorrelations among multiple factors. The theory for this factor has features of a sampling theory, as in Humphreys definition, and features of a compound theory, as stated above. The factor is equivalent to a mixture measure of intelligence. The batteries of tests on which it is based cannot be known to be representative of the "entire repertoire of abilities considered intellectual in nature." The general factor in any one study represents the reliable common variance obtainable with all the tests of a particular battery, but mixtures of batteries of tests vary from one study to another in the same way that mixtures formed by subjective choice of tests vary.

The theory for the factors lower than the general factor in hierarchical theories is the same as Gf–Gc theory, which in turn derives from the theory of the PMA system.

The PMA System

Thurstone's (1938, 1947) research gave us the term *primary mental abilities* (PMA). A metatheory of simple structure guided studies of these abilities. This metatheory specified that abilities largely influence performances on only some tests, not all or even most tests, so there is no general ability pervading performance on all intellectual tests, and most tests indicate (in measurement) only a few abilities, not all abilities. In studies done in the 1930s and 1940s, Thurstone found support for this model in analyses in which no fewer than nine common factors were required to describe most of the reliable individual differences variance obtained with different tests designed to measure important features of intelligence. Examination of the process features of these factors suggested that they indicate basic abilities of Inductive Reasoning (I), Deductive Reasoning (Rs), Practical Problem Reasoning (R), Verbal

Comprehension (V), Associative Short-Term Memory (Ma), Spatial Relations (S), Perceptual Speed (P), Numerical Facility (N), and Word Fluency (Fw).

No fewer than 400 studies (Carroll, 1989) were aimed at replicating the finding of these nine factors and most did. This follow-up research also very much expanded the PMA system. Summaries of this work now indicate (by replicated findings) over 40 factors among tests—factors of approximately the same level of generality as the factors Thurstone identified as *primary* (Eckstrom, French, & Harman, 1979; French, 1951; French, Eckstrom, & Price, 1963; Guilford, 1967; Hakstian & Cattell, 1974; Horn, 1972b). Short descriptions of the best known among these factors are provided in Table 1. The evidence indicates that a model of no fewer than 28 factors is needed to describe the variability measured in tests designed to indicate human intelligence.

As evidence of more and more primary abilities accumulated, it became increasingly clear that a system of 28 to 40 primary abilities is too cumbersome to guide most research. We can't deploy the resources needed to build theories of function for over 28 separate abilities. There must be a rationale for considering smaller numbers of basic cognitive processes. This rationale should accurately take account of the evidence indicating the many factors of the PMA system.

The rationale put forth in accordance with this need for parsimony stemmed from an armchair; that is, the theorist speculated about the organization among abilities and thought up a more parsimonious system. Unfortunately, the rationale often originated only in the armchair and remained there. The evidence of empirical research was not used. In particular, little use was made of research results such as those of the PMA studies, and refutable tests of the system were not carried out. Guilford's work (1967) in developing a theory called the Structure of Intellect Model (SIM) is an important exception to this generalization. Guilford based this theory on the results of PMA studies. This theory is displayed in schematic form in Figure 1.

The basic idea of Guilford's theory is that each of the many abilities humans display can be described as an expression of one of five separate mental operations—Cognition, Memory, Divergent Production, Convergent Production, or Evaluation—operating on one of four separate contents—Figural, Symbolic, Semantic, or Behavioral—to produce one of six separate kinds of products—Units, Classes, Relations, Systems, Transformations, or Implications. As is seen in the figure, such an organizational system implies that there are $5 \times 4 \times 6 = 120$ separate abilities. The 28-plus abilities indicated by PMA research are, according to this theory, particular exemplars of the model. The model is thus parsimonious only in the number of organizing concepts (operations, contents, products) it requires, not in the number of abilities it specifies.

Although there were many studies of the SIM in the 1960s and 1970s, the implications of the system have not been fully explored in research that pro-

**Table 1. First-Order (Primary) Mental Abilities
After Eckstrom, French, & Harman (1979)**

	<i>Guilford Symbol</i>	<i>French Symbol</i>	<i>Repli- cated?</i>
<i>Short-term Apprehension and Retrieval Abilities</i>			
Associative Memory. When presented with one element of previously associated but otherwise unrelated elements, recall the associated element	MSR	Ma	Yes
Span Memory. Immediately recall a set of elements after one presentation.	MSU	Ms	Yes
Meaningful Memory. Immediately recall a set of items that are meaningfully related.	MSR	Mm	?
Chunking Memory. Immediately recall elements by categories into which elements can be classified	MMC		No
Memory For Order. Immediately recall the position of an element within a set of elements	MSS		No
<i>Long-term Storage and Retrieval Abilities</i>			
Associational Fluency. Produce words similar in meaning to a given word.	DMR	Fa	Yes
Expressional Fluency. Produce different ways of saying much the same thing.	DDS	Fe	Yes
Ideational Fluency. Produce ideas about a stated condition or object—e.g., a lady holding a baby.	DMU	Fi	Yes
Word Fluency. Produce words meeting particular structural requirements—e.g., ending with a particular suffix.	DMR	Fw	Yes
Originality. Produce “clever” expressions or interpretations—e.g., titles for a story plot	DMT	O	Yes
Spontaneous Flexibility. Produce diverse functions and classifications—e.g. uses for a pencil.	DMC	Xs	Yes
Delayed Retrieval. Recall material learned hours before			Yes
<i>Visualization and Spatial Orientation Abilities</i>			
Visualization. Mentally manipulate forms to “see” how they would look under altered conditions.	CFT	Vz	Yes
Spatial Orientation. Visually imagine parts out of place and put them in place—solve jigsaw puzzles.	CFS	S	Yes
Speed of Closure. Identify Gestalt when parts of whole are missing.	CFU	Os	Yes
Flexibility of Closure. Find a particular figure embedded within distracting figures	NFT	Cf	Yes
Spatial Planning. Survey a spatial field and find a path through the field—e.g., pencil mazes.	CFI	Ss	Yes
Figural Adaptive Flexibility. Try out in possible arrangements of elements of visual pattern to find one arrangement that satisfies several conditions.	DFT	Xa	Yes
Length Estimation. Estimate lengths or distances between points.		Le	Yes
Figural Fluency. Produce different figures using the lines of a stimulus figure.	DFI		No
Seeing Illusions. Report illusions of such tests as Muller-Lyer, Sanders, & Poggendorff	DFS		No

Table 1 (continued)

	Guilford Symbol	French Symbol	Repli- cated?
<i>Abilities of listening and hearing</i>			
Listening Verbal Comprehension. Show understanding of oral communications.			No
Temporal Tracking. Demonstrate understanding of sequence of auditory information—e.g., reorder a set of tones			No
Auditory Relations. Show understanding of relations among tones—e.g., identify separate notes of a chord.			No
Discriminate Patterns of Sounds. Show awareness of differences in different arrangements of tones.			No
Judging Rhythms. Identify and continue a beat.			No
Auditory Span Memory. Immediately recall a set of notes played once.			No
Perception of Distorted Speech. Demonstrate comprehension of language that has been distorted in several ways.			No
<i>Acculturational Knowledge Abilities</i>			
Verbal Comprehension. Demonstrate understanding of words, sentences and paragraphs.	CMU	V	Yes
Sensitivity to Problems. Suggest ways to deal with problems—e.g., improvements for a toaster.	EMI	Sep	Yes
Syllogistic Reasoning. Given stated premises draw logically permissible conclusions even when these are nonsensical.	EMR	Rs	Yes
Number Facility. Do basic operations of arithmetic quickly and accurately.	NSI	N	Yes
Verbal Closure. Show comprehension of words and sentences when parts are omitted			No
Estimation. Use incomplete information estimate what is required for problem solution.	CMI		No
Behavioral Relations. Judge interaction between people to estimate how one feels about a situation.	CBI		No
Semantic Relations: Esoteric Concepts. Demonstrate awareness of analogic relationships among abstruse bits of information	CMR	NO	MR
Mechanical Knowledge. Information about industrial arts—mechanics, electricity, etc.		Mk	?
General Information: Science, Humanities, Social Sciences, Business		Vi	
<i>Abilities of Reasoning under novel conditions</i>			
Induction. Indicate a principle of relationships among elements.	NSR	I	Yes
General Reasoning. Find solutions for problems having an algebraic quality.	CMS	R	Yes
Figural Relations. Demonstrate awareness of Relationships among figures.	CFR		?
Semantic Relations: Common Concepts. Demonstrate awareness of relationships among common pieces of information.	IMR		

Table 1 (continued)

	Guilford Symbol	French Symbol	Repl-icated?
Symbolic Classifications. Show which symbol does not belong in a class of several symbols.	CSC		No
Concept Formation. Given several examples of a concept, identify new instances.	CFC		No
<i>Speed of Thinking abilities</i>			
Perceptual Speed. Under highly speeded conditions, distinguish similar visual patterns and find instances of a particular pattern.	ESU	P	Yes
Correct Decision Speed. Speed of finding correct answers to intellectual problems of intermediate difficulty.			No
Writing and Printing Speed. As quickly as possible, copy printed or cursive letters or words.			No

vides a sound basis for rejecting or retaining most of its basic hypotheses. The studies designed to test the theory often lacked objectivity (Carroll, 1972; Horn, 1967; Horn & Knapp, 1973, 1974; Humphreys, 1962; Undheim & Horn, 1977). The results overall add up to suggest that some of the three-way combinations of the Guilford system represent empirically based distinctions between established primary mental abilities, but many are only logical indications of ways to construct tests, not indications of distinct human abilities. As

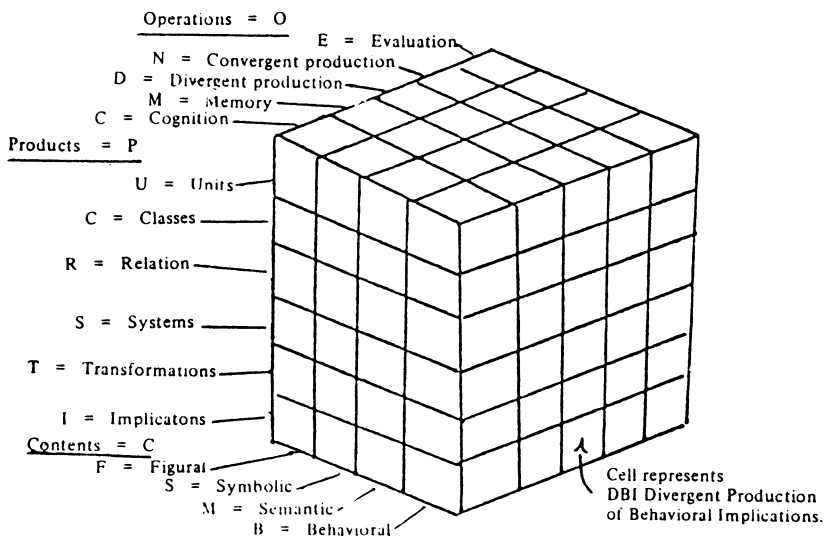


Figure 1. Schematic representation of Guilford's structure of intellect model.

Humphreys (1962, p. 478) pointed out, the facets of the Guilford system “are not psychological as defined. They should be useful to the test constructor, [but] they do not need to make a behavioral difference.” The facet of *products*, particularly, does not appear to correspond to systematic structural or developmental features of human behavior. On the other hand, the facets of *operations*, particularly, and *contents* work better: They point to functions of reasoning, (Cognition), short-term apprehension (Memory), retrieval (Divergent Production), and visualization (Figural Content). These are among the functions of modern-day theory, in particular Gf–Gc theory, which will be considered next.

Gf–Gc Structural Theory

This is a hierarchical theory, derived from PMA theory, that incorporates features of Guilford’s theory. Results from many, many factor analytic studies done in this century add up to suggest the paramorphic organization that is this Gf–Gc theory (Cattell, 1957, 1971; Horn, 1965, 1968, 1988; Horn & Cattell, 1966). Such results indicate that if, in a substantial and heterogeneous sample of people, a researcher measures a variety of cognitive capabilities sampled to represent a simple structure of the abilities specified in the theory, the results from testing to see if the model fits the intercorrelations will indicate that, indeed, such a model does represent the data. It means, too, that many, many other seemingly plausible models do not provide good approximations to the data.

Part of the evidence for the Gf–Gc theory is based on factor analyses of abilities representing the PMA system—that is, structural evidence—but the evidence that mainly distinguishes it from other theories derives from studies of how abilities are affected by brain damage and development, particularly in adulthood. At the structural level, the theory is a second-order system for the PMA factors; that is, a system of factors among factors. The patterns of intercorrelations that point to the system can be seen in many studies (reviewed in, for example, Carroll, 1989, 1993; Horn, 1968, 1976, 1988, 1989a; Horn & Donaldson, 1980). The major results are exemplified in the early study of Horn and Cattell (1967), several follow-up studies by these two investigators (e.g., Cattell & Horn, 1978; Hakstian & Cattell, 1978; Horn & Bramble, 1967; Horn & Stankov, 1982; Rossman & Horn, 1972) and the recent studies of Carroll (1989), Gustafsson (1984), Undheim (1987), and Woodcock (1990). The recent studies, particularly, indicate the generality of the system. Carroll’s results stem from 461 separate studies done by almost as many investigators. Woodcock’s findings are based on a standardization sample of 6,359 subjects spanning an age range from childhood to elderly. Gustafsson’s sample is Swedish, Undheim’s Norwegian. The results indicate that the PMA system can be organized in terms of nine dimensions that are almost as broad as the sets

of abilities people refer to when they use terms such as *intelligence* or *IQ*. Described in capsule form, these abilities are:

- Fluid Reasoning (Gf), measured in tasks requiring inductive, deductive, conjunctive, and disjunctive reasoning to arrive at understanding relations among stimuli, comprehend implications, and draw inferences.
- Acculturation Knowledge (Gc), measured in tasks indicating breadth and depth of the knowledge of the dominant culture.
- Quantitative Knowledge (Gq), measured in tasks requiring understanding and application of the concepts and skills of mathematics.
- Short-term Apprehension retention (SAR) also called short-term memory (Gsm), measured in a variety of tasks that mainly require one to maintain awareness of, and be able to recall, elements of immediate stimulation; that is, events of the last minute or so.
- Fluency of Retrieval from Long-term Storage (TSR), also called long-term memory (Glr), measured in tasks that indicate consolidation for storage and mainly require retrieval, through association, of information stored minutes, hours, weeks, and years before.
- Visual Processing (Gv), measured in tasks involving visual closure and constancy, and fluency in “imaging” the way objects appear in space as they are rotated and flip-flopped in various ways.
- Auditory Processing (Ga), measured in tasks that involve perception of sound patterns under distraction or distortion, maintaining awareness of order and rhythm among sounds, and comprehending elements of groups of sounds, such as chords and the relations among such groups.
- Processing Speed (Gs), although involved in almost all intellectual tasks (Hertzog, 1989), measured most purely in rapid scanning and responding in intellectually simple tasks (in which almost all people would get the right answer if the task were not highly speeded).
- Correct Decision Speed (CDS), measured in quickness in providing answers in tasks that require one to think.

Almost all of the abilities of IQ tests and neuropsychological batteries of tests are accounted for by these nine abilities. That is to say that although IQ tests and neuropsychological batteries are not necessarily described as involving these abilities, nevertheless that which is reliably measured in such tests is mainly predicted and accounted for by the nine factors of Gf–Gc theory. These nine factors also represent the factors of the PMA system in the sense that in a hierarchical system they are higher order organizations of the lower order PMA organizations (Carroll, 1989; Hakstian & Cattell, 1978; Horn & Cattell, 1967).

The Gf–Gc system differs from the early PMA system of Thurstone primarily in the fact that each Gf–Gc factor is broader than the similar factor of Thurstone’s system; that is, it is comprised of, and represents, many more elementary abilities. For example, Gc includes the primaries Verbal Comprehension

sion, Deductive Reasoning, and Numerical Facility, as well as knowledge measured in achievement batteries (Woodcock & Johnson, 1990). Other Gf–Gc factors similarly involve several PMA factors.

The component abilities of each Gf–Gc factor are different, which indicates breadth, but these abilities are also similar relative to the abilities of other Gf–Gc factors. The similarity—this conjunction—is responsible for the factor and indicates the common processes of the ability.

Identifying each factor and replicating this finding in different studies helps to indicate the distinctiveness of each cognitive function. This distinctiveness is demonstrated, also, by showing that factors are construct independent; that is, showing that the best weighted linear combination of any set of eight of the factors does not fully predict the reliable covariance among the component abilities of the ninth factor. This evidence shows that each factor measures a function that is not measured in the other factors.

Construct validity (Campbell & Fiske, 1959; Cronbach, 1987) is indicated by these kinds of evidence. But more evidence is needed to fill out really adequate construct validity, and, indeed, such has been adduced. This evidence indicates how the different abilities develop and operate in human personality. It shows that the factors predict different features of important criteria, stem from different sets of determinants, relate in different ways to estimates of heritability, and are affected in different ways by influences associated with injuries, child rearing, education, using drugs, and other such practices of lifestyle.

In sum, each Gf–Gc factor is broad enough to represent a concept of intelligence, and each involves abilities that are important in definitions of intelligence, but each is distinct from the others when viewed psychometrically, developmentally, in terms of relations to neurology, in terms of what the ability predicts and what predicts it, and in terms of genetic analyses. Each Gf–Gc factor thus represents a separate concept of intelligence. Gf–Gc theory is in this sense a theory of several intelligences, rather than a theory of intelligence.

QUALIFICATIONS

Gf–Gc theory is dependent on the history of research on which it is founded (Sternberg, 1985b). As Sternberg emphasizes, the factors that emerge to support the structural hypotheses of Gf–Gc theory are largely determined by the tasks that are entered into the factor analyses. Rather than choosing tests in accordance with a priori guidelines, tasks were chosen because it was thought they would work to indicate the abilities specified in the theory. In successive studies over the course of this century, tasks were used that were similar to tasks that had been used before. Tradition, Sternberg points out, does not constitute an a priori, theoretical basis for choosing tasks that measure intelligence.

Although this point is important, one should not neglect the fact that all scientific theory is a product of history and culture (Kuhn, 1970). Indeed, scientific theory is, and should be, based on a tradition of observations (induction), as well as on deductions from a consideration of this evidence, the a priori feature of theory that Sternberg emphasizes. The best of scientific theory emerges from an inductive-deductive spiral (Cattell, 1957)—induction followed by deduction, followed by more induction, followed again by deduction, and so on.

Factor-analytic research on human cognitive capabilities has been deductive in two important ways: (a) selections of tasks have been designed to test hypotheses about functions. For example, verbal reasoning tasks were selected in the Horn-Cattell (1967) study to test hypotheses that fluid reasoning (Gf) can involve verbal functions, is not spatial ability and is distinctly different from visual processing (Gv). (b) many of the paradigms, tests and items that have been studied over this century and that form the basis for Gf-Gc theory were based on a priori hypotheses about the nature of intelligence. The results from factor-analytic research are in this sense a distillation of many theories of intelligence. The results confirm and disconfirm many hypotheses.

Factor-analytic evidence provides a basis for structural or organizational features of Gf-Gc theory. Evidence from other kinds of studies (largely group comparison evidence) suggests that the structural organization indicated by factor-analytic research is useful, also, for understanding how different abilities relate to age differences and age changes; to variations in central nervous system (CNS) function, to differences between boys and girls, men and women; to differences in education, social and economic factors; and to variables indicating distinctions between genetic and environmental influences (Horn, 1980b, 1989b).

Gf-Gc structural theory thus guides research on function and development and the neurological and genetic correlates of cognitive development and function. We now turn to a consideration of this research.

DEVELOPMENTAL EVIDENCE

The results of Figure 2 illustrate how mixture measures of *g*, coupled with an assumption that the same *g* is measured in different mixtures, can create confusion. From young adulthood to old age there is (on the average across many individuals) monotonic decrease in some intellectual abilities—namely Gf, Gs, and SAR. Over much of this period, in the same samples of individuals, there are increases in other abilities, in this case Gc and TSR. In very old age, these abilities, too, may decline (Schaie & Baltes, 1977).

Thus, the second-order abilities relate quite differently to age over the adult years. Yet these abilities are parts of what is thrown together in mixture measures of what is called IQ, or *g*, or general intelligence. If such a mixture mea-

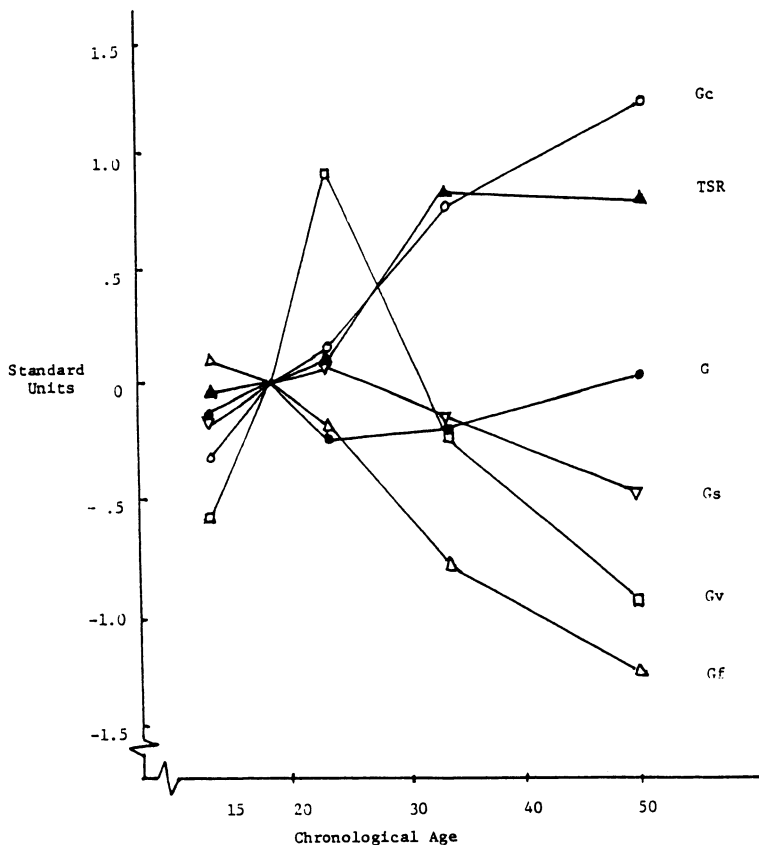


Figure 2. Adulthood age differences in dimensions of human intellect.

sure happens to be made up primarily from abilities that decline with age, then investigators using that measure can argue that intelligence declines with age in adulthood. If most of a mixture measure is made up of abilities for which there is aging increase, those who use that measure can argue that intelligence increases with age in adulthood. If the two kinds of abilities are about equally weighted in a mixture, then the pronouncement can be that intelligence reaches a plateau of growth in adulthood.

Several variations on these themes have been played in the published literature, with resulting controversy and effort to explain the so-called "contradictory" results. For example, many pages in the literature of adult development have been devoted to explanations for a belief that cross-sectional studies show aging decrease in intelligence while longitudinal results do not. This characterization of results is wrong on several counts (Horn & Donaldson, 1980), but one important count is that the apparent contradiction is created from the use

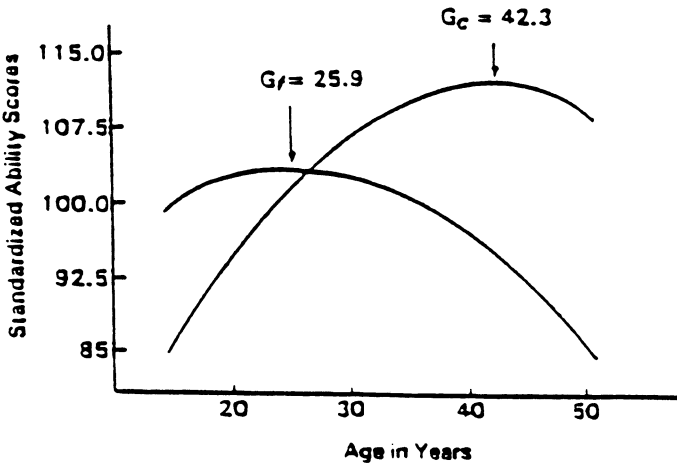


Figure 3. Quadratic factor score curves predicted from model.

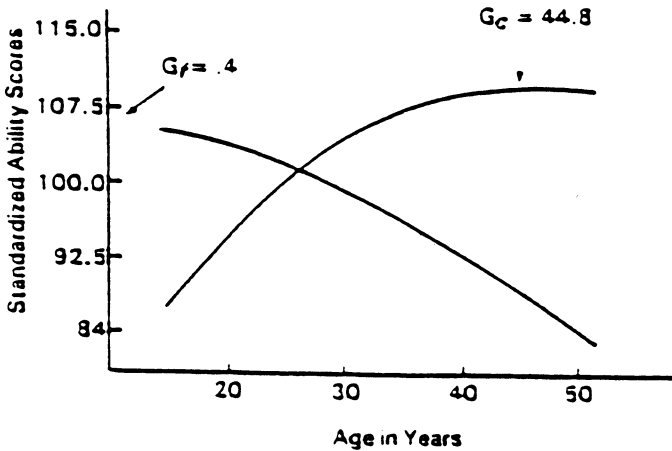


Figure 4. Quadratic factor score curves predicted from education-partialled model.

of mixture measures that were different in longitudinal studies from the mixture measures of cross-sectional studies.

Figures 3 and 4 provide more precise indications than Figure 2 of the age relations for G_f and G_c . The results of these figures are based on structural equation modeling of first-order abilities selected to provide over determination (Thurstone, 1947) of five second-order abilities— G_c , G_f , TSR , G_v , and G_s (McArdle & Horn, 1981). The factor pattern for this model is shown in Table

Table 2. An Exactly Identified Factor Model for G_f, G_c, TSR, G_v & G_s

	<i>Factor Pattern Loadings</i>					
	1	2	3	4	5	<i>U</i> ²
I	.37*		-.00	.37*	-.05	.45*
L	.75*	.19*			-.08	.36*
Ma	.32*		.34*			.71*
V		1.08*		-.09		-.09
Mc		.22*		.34*	-.18*	.32*
EMS	.04	.36*	.04		.23*	.78*
Fa	-.19*	.29	.93*	-.46	-.15	.10*
Fi		.17	.85*	-.34		.43*
CMR	.30*	.60*	.21*		-.19*	.31*
Vz	.18		.08	.63*		.44*
S	.19	-.19*	-.15	.71*	.19*	.33*
Cf	.25	-.01		.52*	.07	.59*
Sc			.41	-.08	.51*	.46*
P	.31*	.03	.07	.13	.46*	.31*
Wf	-.10	-.09	.31	-.07	.33*	.33*
<i>Factor Intercorrelations</i>						
	1					
	.25*	1				
	.37*	.37*	1			
	.29	.45*	.59*	1		
	.28	.02	.36	.04	1	

Notes: *denotes a parameter which is at least twice as big as its standard error. Spaces indicate fixed at zero.

$\chi^2 = 60$, $df = 40$, $z = 2.0$.

2, along with a chi-square, degrees of freedom, and standard errors, all of which help indicate goodness of fit of the model to data.

Figure 5 depicts analyses that provide the basis for the aging curves shown in Figures 3 and 4. In the analyses represented in Figure 5 a linear, quadratic and cubic effect for age, as well as a linear effect for education, was modeled for each of the second-order abilities. The quadratic term (but not the cubic term) is needed to represent the age differences for both G_f and G_c . This is true whether or not variations in education are taken into account, but the effect is most pronounced when the somewhat different educational levels of people of different ages are not controlled (the plot in Figure 3). This plot indicates that the quadratic curve for G_c turns downward in old age when education is not controlled. The curve is at plateau when education is controlled. This suggests that a decline of G_c in old age, an effect that Schaie and Baltes (1977) have been persistent in bringing our attention, may be a cohort effect associated with changes in amount of education throughout much of this century. The quadratic effect at the young end of the curve for G_f , on other hand,

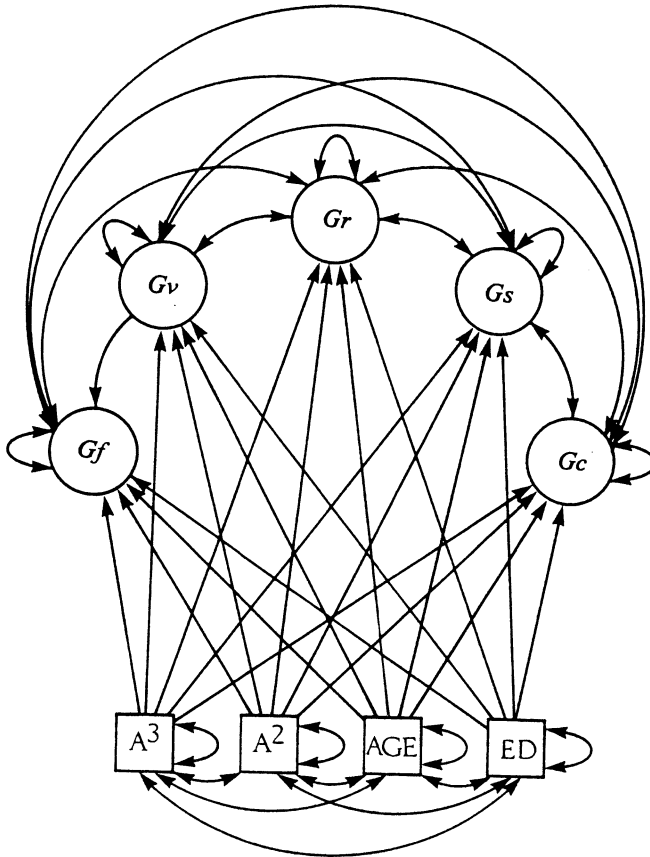


Figure 5. Nonlinear age and education growth curves for factors.

might be a reflection of some of the factors associated with the decline of SAT in samples of young people over the last 10–15 years (Horn, 1978).

It has been argued (Baltes & Schaie, 1974; Schaie, 1973; Schaie & Baltes, 1977) that these results represent only, or mainly, cohort differences, because they are based on cross-sectional data gathering, and that only longitudinal data gathering and mixed designs (Schaie, 1973) can provide dependable evidence for inferences about development. This argument is not valid (Horn & Donaldson, 1976, 1977, 1980; Horn & McArdle, 1980). In fact, each of the methods of data gathering (cross-sectional, longitudinal, and mixtures of the two) has its own particular strengths and weaknesses, and no one of them alone is a royal road to truth. Results obtained with the different methods should be compared and interpreted carefully to derive the unique bits of truth

that can be derived from application of the methods. Figures 3 and 4 of the McArdle-Horn study (and basically even the earlier, cruder Figure 2) correctly depict the results of longitudinal, cross-sectional, and mixed-design studies.

McArdle's (1984) study illustrates a good way in which the results from longitudinal and cross-sectional data can be compared. McArdle brought together data from 58 two-occasion repeated-measures longitudinal studies of the Verbal (V) and Performance (P) measures obtained with the Wechsler Adult Intelligence Scales (WAIS): the V and P measures are similar to Gc and Gf respectively (Matarazzo, 1972). The time between the two testings ranged from 1 hour to 22 years. The sample sizes of the separate comparisons ranged from 1 to 300. The averages for each comparison are depicted in Figure 6. Each of the two ends of a line in this figure is a plot of a point representing the age (along the X-axis) and average score (along the Y-axis). The line between the two endpoints indicates the time between testings, expressed in years of age, and the rise or fall of the average (over people) of a score obtained as the percent of correct answers (rather than the number of correct answers). Lines for which the slope is positive depict increased performance over age (time); lines for which the slope is negative indicate decline of performance with age. The figure on the left is for the V measure; the one on the right is for the P measure.

A first look at Figure 6 might suggest that there are no dependable aging trends in these longitudinal data. The lines seem to go up and down in a fortuitous manner. There is considerable order in these data, however, as McArdle was able to show using mathematical modeling techniques. With these techniques he held constant the test-retest effect over different amounts of time between testings while looking at the age change per year and, similarly, he held constant the age change while looking at time delay between testings.

The results of these analyses are consistent with the much simpler portrayal of results in Figure 2. When other factors are held constant, the aging slope for V changes from positive to negative over the entire age continuum, with most of the consistent part of the negative trend coming in the period beyond age 60. The comparable slope for P is persistently and strongly negative throughout the age period. A structural equation model representing these events provides a reasonable fit to the data. If V and P are combined in a single factor model (to represent the idea of *g*), the resulting model provides a poor fit to the data.

The cross-sectional results shown in Figure 7 illustrate aging trends for V and P that can help us to better understand McArdle's longitudinal results. The findings of Figure 7 also suggest an interesting perspective for the results of Figures 2, 3, and 4 showing individual differences associated with aging. The middle curve in each of the plots in Figure 7 depicts the average at each age; the two curves on either side of the middle curve represent one standard deviation from the average at each age. It can be seen that these cross-sectional

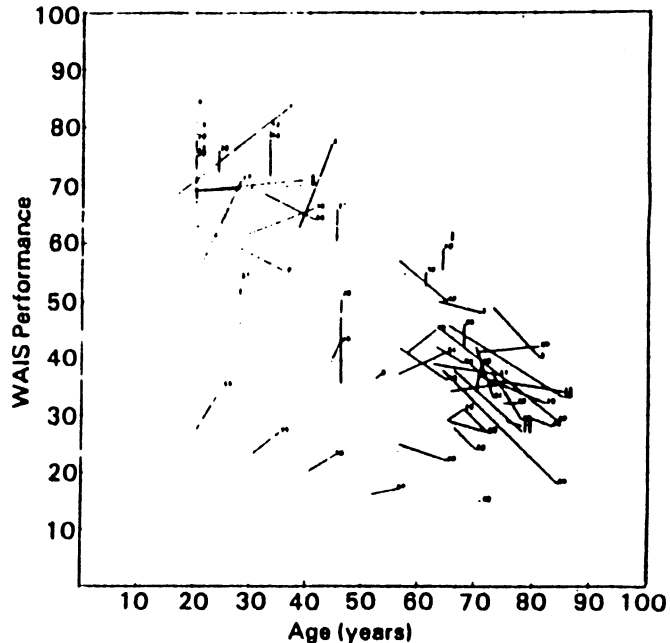
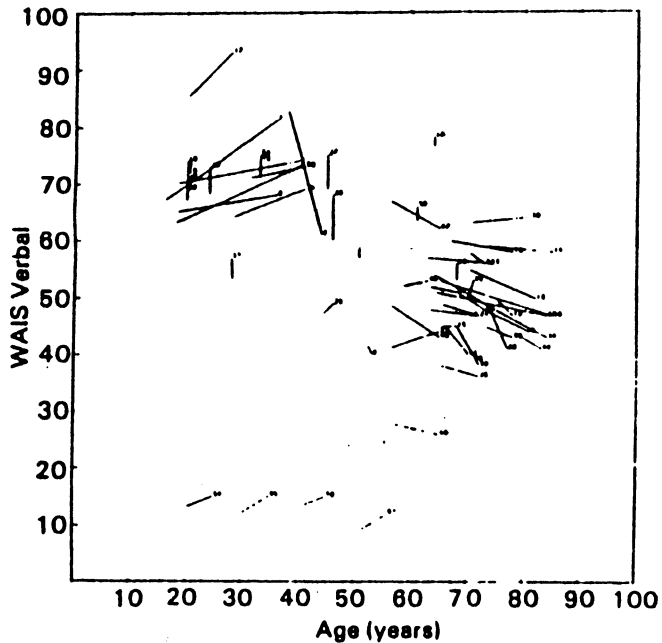


Figure 6. Plot of repeated-measures WAIS studies group means ($N = 58$; verbal and performance).

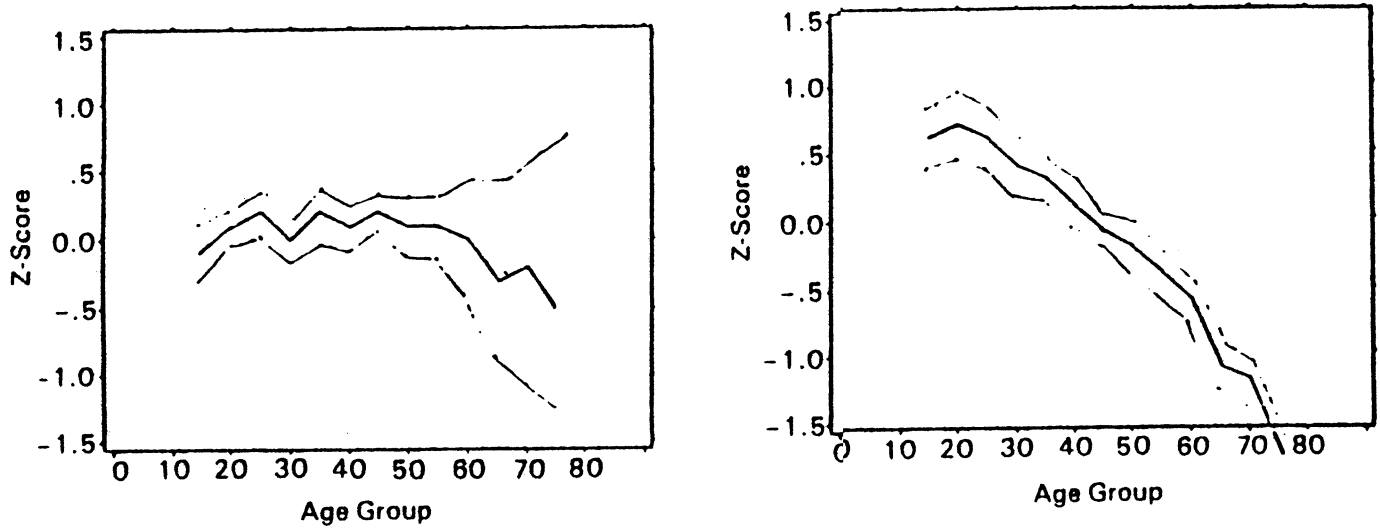


Figure 7. Multiple group factor means and variances from WAIS as a function of age groups in census sample ($N = 2977$).

results are consistent with those just illustrated for the longitudinal data: the slope for the curve of the averages for P is persistently negative; the slope for the curve of the averages for V is near zero or positive in the early years of adulthood, but negative for the most advanced ages (and a quadratic component is needed to well define the curve). But the most remarkable features of these findings are the variances and the differences between the variances for V and P. The variances are similar at each age for P, but they increase notably from ages 40 to 75 years for V.

Because the variance for V is so large at the older ages, a nonrepresentative sample could rather easily contain an overrepresentation of people on either the upper or the lower side of the trend line for the averages shown in Figure 7. If the sampling were on the upper side, the results would suggest that Gc increases with age, whereas if the sampling tended toward the lower side of the trend line, the results would suggest a decrease of Gc in older ages. There is need for further study of these possibilities.

At a more concrete level the results for V in Figure 7 indicate that a considerable subsample of quite elderly individuals perform on the tests of V at a level that is substantially above the level for substantial numbers of younger individuals. This agrees with the notion that some elderly individuals are smarter than many younger individuals. But the results also indicate that a considerable subsample of elderly individuals perform below the average of younger individuals on measures of V. This agrees with the notion that some elderly are not as smart as younger people.

The small and similar variances around the averages for P in Figure 7 suggest that there are not many notable exceptions to the trend showing decline with age in this form of intelligence.

The age-related patterns of change on WAIS measures of P and V appear to be similar for males and females (Kaufman, Kaufman-Packer, McLean, & Reynolds, 1991). Controlling for education differences, no significant age \times sex interaction effects were found for individual differences of verbal or performance scores. Schaie's (1983) hypothesis that women tend to decline first on fluid abilities, whereas men tend to decline first on crystallized abilities, was not supported by this study.

It needs to be kept in mind that although V and P are similar to Gc and Gf, they are not necessarily the best indicators of these concepts. Measures of Gc in the studies of Horn and his colleagues have been better designed than V to represent "practical" knowledge. For example, the general information test of the Horn-Cattell (1966) study sampled knowledge that one could acquire from the everyday living of reading newspapers, watching TV, holding down a job and planning for the future. The comparable tests of the WAIS, on the other hand, more nearly indicate "school learning" rather than the learning of everyday living.

The age differences in averages and variances for Gf and Gc (P and V) do not appear to be due to structural differences: the structural distinction between Gf and Gc (V and P) has been found to be stable across age cohorts (Hayslip & Brookshire, 1985; Cunningham & Birren, 1980). The differences in averages and variances must be understood in terms of different determinants, some of which we will consider next.

ANALYSES OF AGING DECLINE

The decline of Gf over the "vital years of adulthood" is interesting and depressing. The finding of such decline virtually forces the question: "Why?"

There has been widespread belief that the decline is mainly due to aging loss of elementary, relatively unimportant factors such as loss of sensory acuity, increase in carefulness and disinclination to hurry. The evidence does little to support such beliefs.

As concerns disinclination to hurry, there is evidence that with increase in age there is corresponding increase in slowness and carefulness (Birren, 1965, 1974; Horn & Bramble, 1972), but slowness and carefulness appear to stem from recognition of loss of Gf. A series of results lead to this conclusion.

First, there is evidence (Horn et al., 1981; Stankov & Crawford, 1993) of very little relationship (positive but only about $r = .20$) between the speed in getting correct answers to problems of nontrivial difficulty (correct decision speed, CDS) and the level of difficulty of the problems solved. Second, the evidence indicates very little relation between CDS and speed of searching for a particular symbol or deciding whether two stimuli are the same or different (inspection speediness, Gs). Third, and most surprising, the research results indicate that Gs, but not CDS, relates to the aging decline of Gf. These results are summarized in Figures 8 and 9.

The rationale for the analyses of these figures can be illustrated by first considering the case for a claim that the aging decline of Gf is due to loss of a simple function such as CDS (correct decision speed in intellectual problems). It can be seen in Figure 8 that CDS does indeed decline with age. This establishes the prima-facie case: because there is decline in CDS and this is about the same as the decline for Gf, performances on Gf tasks could involve CDS and, therefore, decline of the Gf could result from decline of CDS. This is the nature of most of the evidence arguing that loss of sensory functions and speediness are the factors that produce what is (falsely) interpreted as aging loss of intelligence. Missing in such reasoning, however, is a demonstration that the loss of the simple function is, indeed, associated with the loss of Gf. When this missing link is introduced by controlling for the part of Gf decline that can be accounted for with measure of the elementary process (CDS in this case), the results (illustrated in Figure 9) provide no support for the hypothesis:

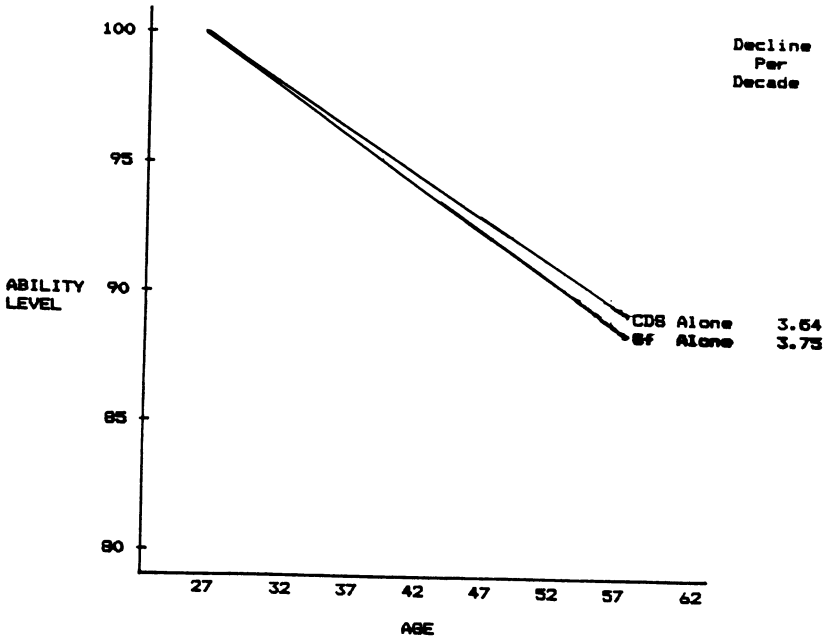


Figure 8. Aging decline of CDS and Gf.

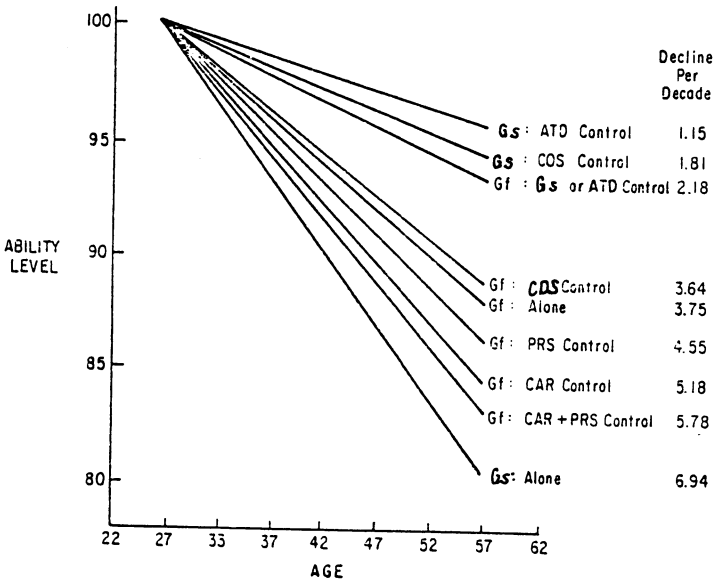


Figure 9. Aging decline of Gf after control for component processes.

the loss of the CDS speediness that accompanies aging does not predict the loss of Gf capabilities. Control for decline of CDS in the decline of Gf does not bring about a significant change: the change from 3.75 to 3.64 units of decline per decade is not significant. This same kind of result has been obtained with measures of auditory and visual detectors and acuities.

Control for inspection speediness, Gs, which also seems to be a rather simple function, does account for some of the aging decline of Gf. The change from 3.75 to 2.15 units of decline per decade is significant. This would seem to support the hypothesis that an aging increase in disinclination to hurry is responsible for the observed Gf decline with age. But there is more to the story. As is seen in Figure 9, most of the aging decline of Gs itself is accounted for by control of a measures of close concentration (COS) and dividing attention (ATD), neither of which involve any speediness: indeed, COS is measured by calibrating how slow one can trace a figure. When COS and ATD are controlled in Gf, there is control also for the part of Gf decline that is associated with Gs. The part of aging decline in Gf that is associated with simple speediness is eliminated by decline of capacities for focusing concentration (COS) and dividing attention (ATD).

Figure 9 also illustrates results suggesting that the decline of Gf does not result because older adults are more careful and/or persistent than younger adults; to the contrary. Older adults work longer than younger adults before abandoning a difficult problem—the PRS (persistence) variable of Figure 9. Older adults also give fewer incorrect answers to problems of nontrivial difficulty—the CAR (carefulness) variable of the figure. CAR and PRS are reflected in slowness of performance in timed tests, particularly if the respondent is given no opportunity to provide an “abandon” response when no solution is found for a problem. But when carefulness and persistence are controlled in the decline of Gf, the decline is not reduced: it is increased! This is seen in the increase in negative slope for curves in which CAR and PRS are controlled. Such findings indicate that carefulness and persistence are qualities that enable older adults to perform better on untimed Gf tasks than they would perform if these qualities were not allowed to operate. When advantages associated with carefulness and persistence are removed by statistical control, there is significant increase in the aging decline of Gf.

Table 3 contains a summary of results from analyses in which several different sets of variables were controlled in studying the aging decline of Gf. One important conclusion derived from these analyses is that different sets of three or four control variables produce essentially the same result; that is, they account for the same amount of aging loss of Gf. This indicates that many ostensibly different variables measure the same basic intellectual processes. For example, although measures of inspection speediness are operationally independent of measures of concentration on slowness and short-term memory, these variables involve a common process (a form of attentiveness) that is

Table 3. Summary of Processes Involved in the Decline of Fluid Abilities Over the "Vital Years" of Adulthood

Choose your favorite three (for any three will do the job of all under cross-validation).	
PROCESS	SYMBOL
1. Concentration: Maintaining close attention, as in doing a task <i>very slowly</i> .	COS
2. Encoding Organization: Classifying incoming information in ways that facilitate subsequent recall.	EOG
3. Incidental Memory: Remembering small things for minutes-hours—i.e. things that would seem to be insignificant.	ICM
4. Eschewing Attentional Irrelevancies: Not attending to what has proved to be irrelevant.	EIR
5. Dividing Attention: Attending to other things while remembering a given thing.	ATD
6. Working Memory: Holding several distinct ideas in mind at once.	MSB
7. Speediness: Speed in "seeing" and "marking" and "comparing".	SPD
8. Hypothesizing: Forming ideas about what is likely.	HYP

From Horn, Donaldson, & Engstrom, *Research on Aging*, 1981, reprinted by permission.

implicated in Gf decline: they do not carry entirely independent variance in accounting for the aging loss of Gf. The same can be said for several other combinations of the variables shown in the table. There are several ways to describe the three or four basic processes that are implicated in aging losses of intellectual abilities (Horn, 1982, 1985a; Horn et al., 1981; Stankov, 1988).

No combination of the variables of Table 3 will account for all the observed decline of Gf. The precise proportion of the decline that is accounted for by different variables varies with the reliabilities of the measures and the extent of the variability in the subject sample, but, roughly, only about one-half of the aging loss of Gf can be reliably accounted for with variables of the kind that are illustrated in Table 3.

GENETIC EVIDENCE AND EARLY DEVELOPMENT

Studies of heritability (Loehlin, Lindsey, & Spuhler, 1975; Plomin, DeFries, & McClearn, 1980; Plomin & Loehlin, 1989; Chipuer, Rovine, & Plomin, 1990) can be useful in indicating how individual differences come about, although it is impossible to fully separate heredity and environment; so the results cannot be interpreted unequivocally. In one kind of heritability research, measures are obtained from samples of identical twins and fraternal twins. Identical twins have the same genetic structure; statistically the expectation is that fraternal

twins share 50% of their genes, the same sharing as for ordinary siblings. In analysis based on these genetic facts the within-pair and between-pair variances and covariances for samples of fraternal and identical twins are partitioned into estimates of the proportions associated with the genetic differences, and lack of differences, of the twin samples. The resulting estimate of heritability depends very much on the reliable variability of the measures and the variability of environmental opportunities within the society where the estimates are obtained (e.g., Herrnstein, 1973; Schonemann, 1990). Such evidence thus needs to be interpreted cautiously.

Most research on the heritability of cognitive capabilities has centered on general intelligence, but Cattell (1940, 1957) put forth hypotheses about Gf and Gc that have been researched. Cattell's first hypothesis stipulated that measured Gf reflects primarily, genetic influences; his second hypothesis stipulated that Gc reflects mainly, environmental influences. He reasoned that Gc stems from Gf; it becomes independent from Gf as individual differences in environmental influences accumulate through childhood (Horn & Cattell, 1966). It follows from this position that in measurement in the earliest period of life one can expect to find virtually no distinction between Gf and Gc, because there would have been few individual differences in environmental influences and little time for such influences to operate, but as development proceeds beyond the earliest years, the distinction between Gf and Gc becomes clearer and clearer. It follows, too, that other things being equal (e.g., the reliabilities), the heritability of Gf should be larger than the heritability of Gc. What does the evidence indicate about these matters?

First, the evidence generally does not support the idea that the distinction between major dimensions of ability will not be seen in early childhood. Very few distinctions in cognitive capabilities can be measured in infants, but Gf, Gc, SAR, and TSR functions can be seen in samples of children as young as four years of age (Horn, 1986; Stankov, 1978). The results of Table 4 illustrates that this distinction is clear at seven years of age.

It is possible that results such as are shown in Table 4 represent individual differences in the environmental influences that operate during the earliest periods of development. In the sample on which these results are based the children had been hospitalized in a neonatal intensive care unit at the time of birth. Children treated in such units differ in respect to medical conditions that necessitate treatment, and they get quite different treatments during hospitalization. Such environmental differences might bring about distinctions that would not be shown in normal development, although this seems unlikely. Further research on this question is needed.

The distinctions indicated in Table 4 could be brought about (in part) by genetic factors. Different bits of evidence are consistent with this hypothesis, and question Cattell's two hypotheses. One line of evidence suggests that the heritability of Gf is not larger than the heritability of Gc. A summary of results

Table 4. Reference Vector Structure Coefficients (Correlations) and Primary Factor Intercorrelations for a Promax Solution based on *n* = 154 7-Year-Olds

<i>Variable Symbols</i>	<i>Descriptions of Variables</i>	<i>Factor Symbols</i>						
		<i>Gc</i>	<i>Gf</i>	<i>TSR</i>	<i>SAR</i>	<i>Gv</i>	<i>Ga</i>	
WEV	Identifying Synonyms	66						
WSP	Identifying Conventional Spelling	65						
KRW	Read Words	53		26				
KCR	Answer Questions about Reading	48		37				
MIN	Answer Questions about Numbers	34	40					
WAR	Show Arithmetic Skills	30	42					
KPS	Put Pictures in Series	51				25		
MCS	Sort Counts by Number	39						
MCB	Group Concepts	36	25					
KMT	Complete Matrix Analogies	33				26		
MVM	Repeat Spoken Words			30	28		30	
KRI	Retrieve Word for Concept		29	55				
KAR	Retrieve Numbers			34				
MVF	Retrieve Words of Categories			32			31	
MWK	Name Pictures	29		38				
KMP	Name Places & People		32			25		
KMN	Repeat Spoken Numbers				64			
KMW	Touch Silhouettes in Order Named				53			
KHM	Repeat (Imitate) Hand Movements				29			
MNB	Repeat Numbers in Reverse Order		28		27			
KTI	Assemble Parts to Match Model					53		
KGC	Gestalt Closure					50		
MDD	Copy Figures					40		
MDC	Draw a Child		36			40		
MVM1	Repeat Spoken Syllables						53	
MVM2	Repeat Spoken Sentences						52	
MD	Follow Spoken Directions						32	
Factor Intercorrelations		Gc	100	19	32	34	28	03
		Gf	19	100	02	30	28	30
		TSR	32	02	100	39	13	05
		SAR	34	30	39	100	31	15
		Gv	28	28	13	31	100	05
		Ga	03	30	05	15	01	100

indicating this outcome is provided in Table 5. This evidence is consistent with an hypothesis that Gf and Gc stem from different genetic determinants, the effects of which can be seen early in development.

Table 5 is based on Nichol's (1978) collation of results from studies of twins. The difference between the intraclass correlations between ability measures for monozygotic and dizygotic twins provides Falconer's (1960) estimate of broad-sense heritability. It can be seen in the table that this heritability is virtually the same for Gc abilities as for Gf abilities. Results reviewed by

Table 5. Primary Ability Average Intraclass Correlations Within Samples of Fraternal and Identical Twins (After Nichols, 1978)

<i>Crystallized Markers</i>	<i>MZ</i>	<i>DZ</i>	<i>Dif.</i>	<i>2 Dif.*</i>
Knowledge: Social Studies	.83	.57	.26	.52
Natural Sciences	.80	.64	.16	.32
V: Verbal Comprehension	.82	.61	.21	.42
Vocabulary	.84	.60	.24	.48
N: Number Facility	.80	.60	.20	.40
Fw: Word Fluency	.65	.51	.14	.28
Fe: Expressional Fluency	.60	.49	.11	.22
Averages	.76	.58	.18	.37
<i>Fluid Markers</i>				
I: Inductive Reasoning	.70	.55	.15	.30
S: Spatial Reasoning	.64	.40	.24	.48
P: Perceptual Speed	.70	.53	.17	.34
N: Number Facility	.80	.60	.20	.40
Ma: Associative Memory	.53	.39	.14	.28
Averages	.67	.49	.18	.36

*Falconer's estimate of broad-sense heritability.

Plomin et al. (1980) and DeFries, Kuse, and Vandenberg (1979) lead to the same conclusion. The average of the heritabilities for abilities that define Gf is no larger than the corresponding average for the abilities of Gc.

Other results support this conclusion (McArdle, Goldsmith, & Horn, 1981). In this work a goodly sample of abilities was selected from the Thurstone, Thurstone, and Standskov (1955) study of twins. The model fitted to partition the variability of the different abilities into parts associated with genetic differences between fraternal and identical twins is for the following four variance-covariance matrices:

- MZB, the between-family variance-covariance matrix for the sample of monozygotic twins ($N = 48$ pairs). To obtain the coefficients of this matrix the measures of the two twins in respect to each variable are added together to provide one measure per family for each variable. The variances and covariances across families for these family measures are the elements of MZB.
- DZB, the between-family variance-covariance matrix for the sample of dizygotic twins ($N = 53$ pairs). This matrix is formed in the same way as MZB; the only difference is that the sample of dizygotic twins is used, rather than the sample of monozygotic twins.
- MZW, the within-family variance-covariance matrix for the sample of monozygotic twins. In this case the difference between the measures for

twins of the same family is the element of the variance for a variable; the covariability of that variability and the same kind of variability for another variable is the covariance of the MZW matrix.

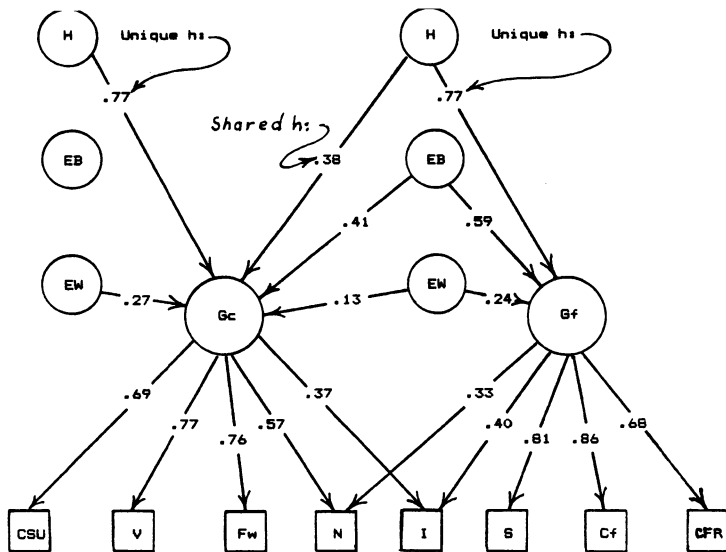
- DZW, the within-family variance–covariance matrix for the sample of dizygotic twins. Again, this is the same as MZW except that it is obtained using the sample of fraternal twins rather than the sample of identical twins.

These matrices represent different proportions of genetic and environmental influences (Eaves & Eysenck, 1977; Eaves, Last, Young, & Martin, 1978). Because identical twins (are assumed to) have the same genetic structure, none of the MZW variance-covariance is due to genetic factors; therefore, all of the variance in the sum of MZB and MZW (which for convenience can be set to 1.0) is partitioned to be associated with MZB. For the sum of DZB and DZW, on the other hand, because only 50% of the genes are on average common to both dizygotic twins, and the square of this 50% is the variance proportion, 25% of the sum-variance can be associated with DZW, which leaves 75% associated with DZB. Similarly, because by definition there can be no between-family environmental variance within families, all of the sum of between and within variance for environmental influences must be associated with the between variance–covariance matrix in both twin samples.

These partitionings of hereditary (H), between-family environmental (EB) and within-family environmental (EW) factors are summarized in the lower part of Figure 10. Also indicated in this figure are specifications for the Gc and Gf dimensions. These were hypothesized to account for the covariance among different variables in the four variance–covariance matrices. The chi-square for the goodness of fit for the model specified in these ways is shown in Figure 10. For reasons discussed rather fully elsewhere in (Horn & McArdle, 1980; Horn et al., 1981), we did not “tidy” the fit. Given these conditions, it can be seen that the specified model provides a reasonable fit to the data.

The variances of all variables—latent and manifest—were standardized to unity in the analyses summarized in the figure. This means that the sum of the squares is unity for the directed relations converging on any particular variable. For example, the sum of the squares is one for the three directed arrows leading into Gf—the .77 from H, the .59 from EB, and the .24 from EW. The square of the .77 for H is the unique broad-sense heritability for Gf. The unique broad-sense heritability for Gc is precisely the same. An hypothesis that Gf heritability is larger than Gc heritability is thus not supported by these findings.

There is nonzero correlation between Gf and Gc. This correlation is modeled as the shared heritability from H to Gc in the figure. This means that in addition to the unique component of genetic variance for Gc, there is a component that stems from Gf. This is consistent with Cattell’s hypothesis that Gc stems from Gf, but the results suggest that this is true of only part of the genet-



The fit is to the following covariance matrices for latent variables
 H=heredity, EB=between environment, EW= within environment

	H	EB	EW
MZB: Monozygotic between families	1	1	1
MZW: Monozygotic within families	0	0	1
DZB: Dizygotic between families	.75	1	1
DZW: Dizygotic within families	.25	0	1

Figure 10. A Genetic Model of Gf and Gc (McArdle, Goldsmith, & Horn, 1981).

ic variance of Gc. That is, part of the heritable variability of Gc can be said to be due to a unique set of genes and part is due to a set of genes that also are part of the heritability of Gf. The shared variability can represent either a fact that the Gf and Gc constructs must share common genetic determiners or a fact that the fallible measures of Gf and Gc are not fully independent in measurement. An intricate design is required to separate these two possible interpretations.

These results of studies of heritability and early childhood distinction between cognitive capabilities thus lead to a conclusion that different intelligences can be distinguished in early childhood, partly because they stem from separate genetic determiners and partly because they stem from separate environmental determiners. It follows logically that the observed distinctions probably mirror distinctions in neurological functions.

HERITABILITY OF GENERAL INTELLIGENCE

Jensen (1973) has likened the heritability IQ (different mixture measures) to the heritability of a polygenetic trait. He has referred to the quasi-normal distribution of IQ measures and the regression of IQ scores for related people—as from parent to child—as evidence in support of this theory. As we have seen, the extant evidence does not support the hypothesis that cognitive capabilities are unitary, but setting that problem aside, there are problems with Jensen's line of reasoning.

The alleles of perhaps several genes, transmitted independently, add up to determine a polygenetic trait. Skin pigmentation represents such a trait: if none of the alleles for such pigmentation are transmitted, the individual will be very white; if one gets all the alleles, pigmentation is very heavy and the person is jet black; the colors of most people are between these extremes, distributed in accordance with a symmetrical binomial, which is an approximation to a normal distribution.

There could be an attribute called intelligence that conforms to this kind of theory, but finding of a symmetrical binomial (normal) distribution for measures of IQ neither supports nor threatens such a theory. Gene determiners produce the normal distribution of the theory, but responses to items used to measure ability produce the distribution of IQ scores. There is no known isomorphism between these two kinds elements—alleles of genes and responses to items—and there are no compelling reasons to suppose that such isomorphism might exist. The influences that determine which items are put into an IQ test are not at all similar to the influences that determine gene selection in reproduction.

Environmental influences, too, can combine independently to produce a normal distribution, but in this case, too, a finding of an approximately normal distribution for measures of IQ neither supports nor refutes a claim that the trait is determined by environmental factors; there is no reason to expect that the items of the IQ test are isomorphic to environmental influences.

The point is that the shape of a distribution of scores on a test provides no evidence that the characteristic the scores are intended to indicate either is or is not determined by environmental or genetic influences. Shape of distribution is irrelevant to the claims.

Similarly irrelevant is the finding of regression to the mean of high or low measures of IQ on one class of people (e.g., children), relative to similar measures on another class of people (e.g., parents). Such regression is no more than a restatement of the fact that the two arrays of measures are less than perfectly correlated—that is, are somewhat independent. The independence could reflect the fact that genes are sorted independently in the two parents that transmit genetic potential to a person, but it can reflect any of many other influences, not the least being that environmental influences affecting a person

are somewhat independent of the comparable influences affecting parents. Regression, as such, is no more an indicator of genetic transmission than it is an indicator of environmental influences. It is not evidence in support of a polygenetic theory of IQ.

This is not to argue that there is no evidence that mixture measures of cognitive capabilities are inherited. There is such evidence. Relevant results include correlations between IQ scores for people of different degrees of biological relationship. These are roughly in order of the extent to which people are genetically related. For identical twins reared together (with the same genetic structure) the correlations are about .8; for fraternal twins (who share genes to the same extent as ordinary siblings, or a child and parent), the correlations are about .6-.7; for ordinary siblings and the parent-child correlations are approximately .5; for half siblings, the correlations are roughly .3-.4; the correlations for first cousins, uncle-nephew, aunt-niece, child-grandparent are lower than this, but larger than for unrelated people raised in the same home (correlations between 0 and .3) or unrelated people randomly assigned to pairs (correlations near 0.0).

Such results are consistent with an hypothesis that some of the individual differences variability collections of cognitive abilities stem from genetic factors. But the order of similarity in genetic relationships is also the order of similarity one can expect for the environments in which people develop. The environments for identical twins are most similar, for fraternal twins next most similar, for ordinary siblings next after this, and for other comparisons, too, the similarities expected for environments decreases monotonically with decrease in similarities in genetic structure.

Genetic and environmental influences thus are confounded. No way is known in research design or analysis to unconfound the influences, and be sure that they are unconfounded in results. This is true for the results summarized in Figure 10 as well as for results for putative measures of general intelligence, but the findings of Figure 10 indicate separate heritabilities for Gf and Gc, which does not support the hypothesis that general intelligence is a polygenetic trait. The shared *h* in the model could indicate common genetic determiners, but could just as well indicate inability to distinguish Gf and Gc in fallible measurement.

All in all the evidence is consistent with an hypothesis that the different mixture measures that are called general intelligence are mixtures of attributes that are inherited separately rather in the way that eye, ear, cheekbone, lip, and nose characteristics of the face are inherited separately. Similarly, structures of the brain are inherited separately and these support cognitive capabilities. The brain structures are separately located, are functionally distinct, and genetically independent.

For example, the norepinephrine system centers around the locus coeruleus, branching largely into the hypothalamus and adjacent areas, is closely associ-

ated with arousal of neurological functions—arousal such as appears to be manifested in Gf (Iverson, 1979; Horn, 1982, 1985a, for review). The dopamine system, on the other hand, centered around the substantia nigra and corpus striatum, is linked to a complex of events associated with such outcomes as Parkinson's disease. The serotonin system, also, has a distinct place of function in the brain and distinct associations in behavior.

Anatomical analyses also indicate distinct functions associated with different sections of the brain. The left hemisphere, for example, is associated with different aspects of intellectual function than is the right hemisphere, and a growing mound of evidence suggests that the top-to-bottom and front-to-back divisions of the brain are even more important indicators of distinct ability functions than is the left-to-right division (Blackwood & Corsellis, 1976; Bourne, Ekstrand, & Dominowski, 1971; Prohovnik, 1980).

The structures and functions of the brain thus are distinct, and derive from different genetic determinants. The different structures and functions have a different role to play in sensation, perception, and learning. Different configurations of these distinct features produce different cognitive capacities, different perceptions, and different ways to process the same information. Just as there are many different configurations of facial features that provide examples of a "beautiful face," so there are many configurations of features of brains that exemplify the "good brain" that underlies good intelligence. The terms intelligence and beauty unite diversity in a single word in colloquial language, but this is not the language of science. The words are not expressions of lawful functions. Studies of how the brain functions, and of different brains, do not support any known theory of general intelligence. To the contrary, this evidence suggests that there should be several intelligences.

SUMMARY AND CONCLUSIONS

Evidence accumulated over the course of this century has made it clear that the phenomena of human intelligence is multidimensional. The evidence of individual differences (structural evidence), of change from infancy to old age (developmental evidence), of relationships to indicators of physiological and neurological (neurocognitive evidence), and of relationship among persons related biologically in different degrees (heritability evidence), indicate that although one word, intelligence, is used to refer to human cognitive capabilities, a single scientific concept does not represent the phenomena. That which is referred to as intelligence is a melange of many different cognitive capabilities.

Common factors among abilities have been found at four levels. The number of factors now established at what is known as the primary mental ability level is somewhat more than 40. These abilities account for redundancy in

measures of many hundreds of tests. But these 40-plus abilities are not entirely independent: They, too, order individuals in somewhat the same way. When the redundancy among these primary abilities is analyzed, evidence of 10 factors is found at what is known as the second-order. These factors also are not entirely independent: when their redundancy is analyzed, two common factors are found. These two factors are positively correlated, which indicates one common factor.

A system for explaining this evidence is the theory of fluid (Gf) and crystallized intelligence (Gc), known simply as Gf–Gc theory. This theory has changed with the accumulation of evidence. At first it was a theory of two intelligences: today, it would be better labeled the theory of many intelligences. The principal concepts derive from structural evidence at the second-order. These concepts are: Fluid Reasoning (Gf), Acculturation Knowledge (Gc), Quantitative Knowledge (Gq), Visual Processing (Gv), Auditory Processing (Ga), Processing Speed (Gs), Correct Decision Speed (CDS), Short-term Apprehension-Retention (SAR), and Fluency of Retrieval from Long-term Storage (TSR). Concepts of auditory and visual sensory detection, carefulness, concentration, attention, incidental memory, speediness and hypothesizing also are used in the theory. Each of the second-order abilities has, on its own, been regarded as equal to, or central to, general intelligence (IQ or *g*). Indeed, Gf and Gc have been discussed as if they are equivalent to each other—as if each indicates the same IQ or *g*, although the evidence indicates that these two factors have quite different construct validities. Short-term Apprehension and Retrieval and Quick thinking, too, have been regarded as the central features of intelligence. No one of these capabilities well represents all the others or is the essence of all the phenomena referred to with the word intelligence. They are better regarded as distinct intelligences.

The abilities of separate intelligences increase with learning, practice and use. Without use (practice) they decline. They decline also with loss of neurological base. Over the ages of childhood the averages for all cognitive abilities increase. As development continues through adulthood, the averages for Gc and TSR continue to increase, but the averages for Gf, SAR, and Gs decrease. The evidence is consistent for both cross-sectional and longitudinal data.

Gf and SAR are said to be vulnerable abilities. Not only are these the abilities that decline first and most with age in adulthood, they are most irreversibly affected by—are most vulnerable to—injuries to the central nervous system (CNS). In contrast to the vulnerable abilities, Gc and TSR are maintained abilities: they do not decline with age over most of adulthood. Also, although these abilities are depressed immediately following by brain damage (such as that produced by stroke) they spring back” to nearly preinjury level in the weeks of the recovery period.

Early studies were directed at establishing the extent to which general intelligence is inherited. The results were confusing, partly because the concept of

general intelligence was poorly defined: measures were mixtures of different intelligences. Cattell put forth hypotheses that helped focus research. He suggested that Gf reflects primarily genetic influences, and Gc reflects mainly environmental influences. The genetic potential of Gf as thought to be "invested" in Gc over the course of development, but Gc was also produced through environmental conditions for which there are individual differences. Therefore, Gf and Gc would not be distinct early in development, but would become distinct as development proceeds into later childhood and adulthood.

Most of the evidence bearing on these hypotheses does not support Cattell's main arguments. In 11 separate studies based on analyses of monozygotic and dizygotic twins the heritability estimates for Gf abilities were not systematically larger than the heritability estimates for Gc abilities. Other results indicate that Gf and Gc are distinguishable in early childhood. The results are consistent with an hypothesis that Gf and Gc stem from different genetic determinants, the effects of which can be seen early in development.

The distinctiveness of separate intelligences thus has been demonstrated by evidence that the different capabilities have different relationships with other variables, develop differently, relate differently to variables of physiological structure and function, and to those of education and genetics. The different capabilities thus appear to stem from different sets of determinants and to be affected in different ways by influences associated with injuries, childrearing, education, and the variety of practices that make up different lifestyles.

We see that some things are known about what is called intelligence, but that much more is not known. This large unknown calls for new scientific research.¹

REFERENCES

- Alexander, W. P. (1935). Intelligence, concrete and abstract. *British Journal of Psychology* (Monograph Suppl. No. 19).
- Baltes, P. B., & Schaie, K. W. (1974). The myth of the twilight years. *Psychology Today*, 40, 35-40.
- Biederman, I. (1992). Human image understanding. In P. Johansen & S. Olsen (Eds.), *Theory and applications of image analyses* (pp. 3-14). Singapore: World Scientific.
- Binet, A., & Simon, T. (1905). Upon the necessity of establishing a scientific diagnosis of inferior states of intelligence. *L'Annee Psychologique*, 11, 163-190; 191-244, 245-366.
- Blackwood, W., & Coresellis, J. A. (Eds.). (1976). *Greenfield's neuropathology*. London: Arnold.

¹We thank Dr. Lazar Stankov for very helpful suggestions he provided in review of this manuscript.

- Bourne, L. R., Ekstrand, B. R., & Dominowski, R. L. (1971). *The psychology of thinking*, Englewood Cliffs, NJ: Prentice-Hall.
- Brigham, C. C. (1932). *The study of error*. New York: College Entrance Examination Board.
- Broadbent, D. E. (1966). The well-ordered mind. *American Educational Research Journal*, 3, 281–295.
- Brown, W. (1933). The mathematical and experimental evidence for the existence of a central intellectual factor. *British Journal of Psychology*, 23, 171–179.
- Burt, C. (1909). Experimental tests of general intelligence. *British Journal of Psychology*, 3, 94–177.
- Burt, C. (1911). Experimental tests of higher mental processes and their relation to general intelligence. *Journal of Experimental Pedagogy*, 1, 93–112.
- Burt, C. (1941). *Factors of the mind*. London: University of London Press.
- Burt, C. (1949a). Subdivided factors. *British Journal of Statistical Psychology*, 2, 41–63.
- Burt, C. (1949b). The structure of the mind: A review of the results of factor analysis. *British Journal of Psychology*, 19, 176–199.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.
- Carroll, J. B. (1972). Stalking the wayward factors. *Contemporary Psychology*, 17, 321–324.
- Carroll, J. B. (1989). Factor analysis since Spearman: Where do we stand? What do we know? In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology: The Minnesota symposium on learning and individual differences* (pp. 43–67). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University Press.
- Cattell, R. B. (1940). A culture-fair intelligence test 1. *Journal of Educational Psychology*, 31, 161–179. (Revised in 1975).
- Cattell, R. B. *Personality and motivation structure and measurement*. New York: World Book.
- Cattell, R. B. (1971). *Abilities: Their structure, growth and action*. Boston: Houghton Mifflin.
- Cattell, R. B., & Horn, J. L. (1978). A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *Journal of Educational Measurement*, 15, 139–164.
- Chipuer, H. M., Rovine, M. J., & Plomin, R. (1990). LISREL modeling: Genetic and environmental influences on IQ revisited. *Intelligence*, 14(1), 11–29.
- Commons, M. (1985). *How novelty produced continuity in cognitive development within a domain and accounts for unequal development across domains*. Toronto: Society for Research of Child Development.
- Cox, G. W. (1928). *Mechanical aptitude*. London: Methuen.
- Cronbach, L. J. (1977). *Educational psychology* (3rd ed.). New York: Harper & Row.
- Cronbach, L. J. (1987). Construct validation after thirty years. In R. Linn (Ed.), *Intelligence: Measurement, theory and public policy*. Urbana, IL: University of Illinois Press.
- Cunningham, W. R., & Birren, J. E. (1980). Age changes in the factor structure of intellectual abilities in adulthood and old age. *Educational and Psychological Measurement*, 40, 271–290.

- DeFries, J. C., Kuse, A. R., & Vandenberg, S. G. (1979). Genetic correlations, environmental correlations and behavior. In J. R. Royce & L. P. Mos (Eds.), *Theoretical advances in behavior genetic* (pp. 389–421). Alpen aan den Rijn, Netherlands: Sijthoff Noordhoff International.
- Detterman, D. K. (1989). Editorial: The future of intelligence research. *Intelligence*, *13*, 199–203.
- Eaves, L. J., & Eysenck, H. J. (1977). A genotype-environmental model for psychotism. *Advances in Behavioral Research Theory*, *1*, 5–26.
- Eaves, L. J., Last, K. A., Young, P. A., & Martin, N. G. (1978). Model-fitting approaches to the analysis of human behavior. *Heredity*, *41*, 249–320.
- Eckstrom, R. B., French, J. W., & Harman, M. H. (1979). Cognitive factors: Their identification and replication. *Multivariate Behavioral Research Monographs*, *79*.
- El Koussy, A. A. H. (1935). The visual perception of space. *British Journal of Psychology Monograph Supplement*, *7* (20).
- Eysenck, H. J. (Ed.). (1982). *A model for intelligence*. Berlin: Springer-Verlag.
- Falconer, D. S. (1960). *Introduction to quantitative genetics*. New York: Ronald Press.
- French, J. W. (1951). The description of aptitude and achievement tests in terms of rotated factors. *Psychometric Monographs*, No. 5.
- French, J. W., Eckstrom, R. B., & Price, L. A. (1963). *Manual and kit of reference tests for cognitive factors*. Princeton, NJ: Educational Testing Service.
- Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, *5*, 351–360.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Gustafsson, J. E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, *8*, 179–203.
- Gustafsson, J. E. (1985). Measuring and interpreting g. *The Behavioral and Brain Sciences*, *8*, 231–232.
- Hakstian, A. R., & Cattell, R. B. (1974). The checking of primary ability structure on a broader basis of performances. *British Journal of Educational Psychology*, *44*, 140–154.
- Hakstian, A. R., & Cattell, R. B. (1978). Higher stratum ability structure on a basis of twenty primary abilities. *Journal of Educational Psychology*, *70*, 657–659.
- Hayslip, B., & Brookshire, R. G., (1985). Relationships among abilities in elderly adults: A time lag analysis. *Journal of Gerontology*, *40*(6), 748–750.
- Herrnstein, R. J. (1973). *IQ in the meritocracy*. Boston: Little, Brown.
- Hertzog, C. (1989). Influences of cognitive slowing on age differences. *Developmental Psychology*, *25*, 636–651.
- Holzinger, K. J., & Harman, H. H. (1941). *Factor analysis: A synthesis of factorial methods*. Chicago: University of Chicago Press.
- Horn, J. L. (1965). *Fluid and crystallized intelligence: A factor analytic and developmental study of the structure among primary mental abilities*. Unpublished doctoral dissertation, University of Illinois.
- Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, *75*, 242–259.
- Horn, J. L. (1970). The nature of human intelligence by J. P. Guilford. *Psychometrika*, *35*, 273–277.
- Horn, J. L. (1972a). State, trait and change dimensions of intelligence. *British Journal of Educational Psychology*, *42*, 159–185.

- Horn, J. L. (1972b). The structure of intellect: Primary abilities. In R. M. Dreger (Ed.), *Multivariate personality research* (pp. 451–511). Baton Rouge, LA: Claitor's Publishing Company.
- Horn, J. L. (1976). Human abilities: A review of research and theory in the early 1970s. *Annual Review of Psychology*, 27, 437–485.
- Horn, J. L. (1978). The nature and development of intellectual abilities. In R. T. Osborne, C. E. Noble, & N. Weyl (Eds.), *Human variation: The biopsychology of age, race, and sex* (pp. 107–136). New York: Academic Press.
- Horn, J. L. (1980). Concepts of intellect in relation to learning and adult development. *Intelligence*, 4, 285–317.
- Horn, J. L. (1981). Multiple measurement theory and technology for diagnosing intellectual deficiencies of aging adults. Drugs and methods In *Proceedings of the International Symposium on Experimental and Clinical Methodologies for Study of Acute and Chronic Cerebrovascular Diseases* (pp. 15–48). New York: Pergamon.
- Horn, J. L. (1982). The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood. In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 237–278). New York: Plenum.
- Horn, J. L. (1985). Remodeling old models of intelligence: Gf-Gc theory. In B. B. Wolman (Ed.), *Handbook of intelligence* (pp. 267–300). New York: Wiley.
- Horn, J. L. (1988). Thinking about human abilities. In J. R. Nesselroade (Ed.), *Handbook of multivariate psychology* (pp. 645–385). New York: Academic Press.
- Horn, J. L., & Bramble, W. J. (1967). Second order ability structure revealed in right and wrong scores. *Journal of Educational Psychology*, 58, 115–122.
- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology*, 57, 253–270.
- Horn J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, 26, 107–129.
- Horn, J. L., & Donaldson, G. (1976). On the myth of intellectual decline in adulthood. *American Psychologist*, 31(10), 701–719.
- Horn, J. L., & Donaldson, G. (1977). Faith is not enough: A response to the Baltes-Schaie claim that intelligence does not wane. *American Psychologist*, 32(5), 369–373.
- Horn, J. L., & Donaldson, G. (1980). Cognitive development in adulthood. In O. G. Brim & J. Kagan (Eds.), *Constancy and change in human development* (pp. 445–529). Cambridge, MA: Harvard University Press.
- Horn, J. L., Donaldson, G., & Engstrom, R. (1981). Apprehension, memory, and fluid intelligence decline in adulthood. *Research on Aging*, 3(1), 33–84.
- Horn, J. L., & Knapp, J. R. (1973). On the subjective character of the empirical base of Guilford's Structure-of-Intellect model. *Psychological Bulletin*, 80, 33–43.
- Horn, J. L., & Knapp, J. R. (1974). Thirty wrongs of not make a right: Reply to Guilford. *Psychological Bulletin*, 81, 502–504.
- Horn, J. L., & McArdle, J. J. (1980). Perspectives on mathematical/statistical model building (MASMOB) in research on aging. In L. W. Poon (Ed.), *Aging in the 1980's* (pp 503–541). Washington DC: American Psychological Association.
- Horn, J. L., & Stankov, L. (1982). Auditory and visual factors of intelligence. *Intelligence*, 6, 165–185.

- Hultsch, D. F. (1971). Adult age differences in free classification and free recall. *Developmental Psychology*, 4, 338-342.
- Humphreys, L. G. (1962). The organization of human abilities. *American Psychologist*, 17, 475-483.
- Humphreys, L. G. (1979). The construct of general intelligence. *Intelligence*, 3, 105-120.
- Iverson, L. L. (1979). The chemistry of the brain. *Scientific American*, 241, 134-149.
- Jensen, A. R. (1973). Level I and Level II abilities in three ethnic groups. *American Educational Research Journal*, 10, 263-276.
- Jensen, A. R. (1982). Reaction time and psychometric g. In J. J. Eysenck (Ed.), *A model for intelligence* (pp. 93-132). New York: Springer-Verlag.
- Jensen, A. R. (1984). Test validity: g versus the specificity doctrine. *Journal of Social and Biological Structures*, 7, 93-118.
- Kaufman, A. S., Kaufman-Packer, J. L., McLean, J. E., & Reynolds, C. R. (1991). In the pattern of intellectual growth and decline across the adult life span different for men and women? *Journal of Clinical Psychology*, 47 (6), 801-812.
- Kelley, T. L. (1928). *Crossroads in the mind of man: A study of differentiable mental abilities*. Stanford, CA: Stanford University Press.
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.
- Loehlin, J. C., Lindsey, G., & Spuhler, J. N. (1975). *Race differences in intelligence*. San Francisco: W. H. Freeman.
- Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). Baltimore, MD: Williams and Wilkins.
- McArdle, J. J. (1984, November). *A dynamic and structural equation model of WAIS abilities*. Paper presented at the Annual Meeting of the Society of Multivariate Experimental Psychologists, Evanston, IL.
- McArdle, J. J., Goldsmith, H. H., & Horn, J. L. (1981). Genetic structural equation models of fluid and crystallized intelligence. *Behavior Genetics*, 60, 607.
- Nichols, R. (1978). Twin studies of ability, personality, and interests. *Homo*, 29, 158-173.
- Patterson, D. G., & Elliot, R. N. (1930). *Minnesota mechanical ability tests*. Minneapolis: University of Minnesota Press.
- Patterson, D. G., & Elliot, R. N. (1930). *Minnesota mechanical ability tests*. Minneapolis: University of Minnesota Press.
- Plomin, R., DeFries, J. C., & McClearn, G. E. (1980). *Behavioral genetics*. San Francisco: Freeman.
- Plomin, R., & Loehlin, J. C. (1989). Direct and Indirect IQ Heritability Estimates: A Puzzle. *Behavior Genetics*, 19(3), 331-342.
- Porteus, S. D. (1946). *The Porteus Maze Test: Revised*. New York: Psychological Corporation. (Original work published in 1914)
- Prohovnik, I. (1980). *Mapping brainwork*. Malmö, Sweden: CWK Gleerup.
- Raven, J. C. (1938). *Guide to using progressive matrices*. London: H. K. Lewis. (Revised 1954, 1956).
- Rimoldi, H. J. A. (1948). Study of some factors related to intelligence. *Psychometrika*, 13, 27-46.
- Rossman, B. B., & Horn, J. L. (1972). Cognitive, motivational and temperamental indi-

- cants of creativity and intelligence. *Journal of Educational Measurement*, 9, 265–286.
- Schaie, K. W. (1973). Methodological problems in descriptive developmental research on adulthood and aging. In J. R. Nesselroade & H. W. Reese (Eds.), *Life-span developmental psychology: Developmental issues*. New York: Academic Press.
- Schaie, K. W. (Ed.). (1983). *Longitudinal studies of adult psychological development*. New York: Guilford.
- Schaie, K. W., & Baltes, P. B. (1977). Some faith helps see the forest: A final comment on the Horn and Donaldson myth of the Baltes-Schaie position on adult intelligence. *American Psychologist*, 32, 1118–1120.
- Schonemann, P. H. (1990). New questions about old heritability estimates. *Bulletin of the Psychonomic Science*, 27, 175–178.
- Spearman, C. (1904). "General Intelligence," objectively determined and measured. *American Journal of Psychology*, 15, 210–293.
- Spearman, C. (1927). *The abilities of man: Their nature and measurement*. London: Macmillan.
- Spearman, C. (1937). *Psychology down the ages*. London: Macmillan.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, 74, 498–450.
- Stankov, L. (1978). Fluid and crystallized intelligence and broad perceptual factors among the 11 to 12 year olds. *Journal of Educational Psychology*, 70(3), 324–334.
- Stankov, L. (1988). Aging, intelligence and attention. *Psychology of Aging*, 3(2), 59–74.
- Stankov, L., & Crawford, J. D. (in press). Ingredients of complexity in fluid intelligence. *Learning and Individual Differences*.
- Stern, W. (1914). The psychological methods of testing intelligence (G. M. Whipple, Trans.). *Educational Psychological Monographs* (No. 3). Baltimore, MD: Warwick & York.
- Sternberg, R. J. (1985). General intellectual ability. In R. J. Sternberg (Ed.), *Human abilities*. New York: Freeman.
- Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge: Cambridge University Press.
- Sternberg, R. J., & Detterman, D. K. (Eds.). (1986). *What is intelligence? Contemporary viewpoints on its nature and definition*. Norwood, NJ: Ablex.
- Terman, L. M. (1916). *The measurement of intelligence*. Boston: Houghton Mifflin.
- Thomson, G. A. (1919). On the cause of hierarchical order among correlation coefficients. *Proceedings of the Royal Society, A*, 95, 400–408.
- Thomson, G. A. (1948). *The factorial analysis of human abilities* (3rd ed.). Boston: Houghton-Mifflin.
- Thurstone, L. L. (1938). Primary mental abilities. *Psychometric Monographs* (No. 1). Chicago: University of Chicago Press.
- Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.
- Thurstone, T. G., Thurstone, L. L., & Standskov, H. H. (1955). *Psychological study of twins. Scores of one hundred and twenty-five pairs of twins on fifty-nine tests*. Chapel Hill, NC: The Psychometric Laboratory, University of North Carolina.
- Undheim, J. O. (1987). The hierarchical organization of cognitive abilities: restoring general intelligence through the use of Linear Structural Relations (LISREL). *Multivariate Behavioral Research*, 22, 149–171.

- Undheim, J. O., & Gustafsson, J. E. (1987). The Hierarchical Organization of Cognitive Abilities: Restoring General Intelligence Through the Use Linear Structural Relations (LISREL). *Multivariate Behavioral Research*, 22, 149–171.
- Undheim, J. O., & Horn, J. L. (1977). Critical evaluation of Guilford's structure-of-intellect theory. *Intelligence*, 1, 65–81.
- Willoughby, R. R. (1927). Family similarities in mental-test abilities. *Genetic Psychology Monographs*, 2, 235–277.
- Woodcock, R. W. (1990). Theoretical foundations of the WJ-R measures of cognitive ability. *Journal of Psycho-Educational Assessment*.
- Woodcock, R. W., & Johnson, M. B. (1990). *Woodcock-Johnson Psycho-Educational Battery* (revised). Allen, TX: DLM Teaching Resources.

Chapter 8

*Intelligence from the
Standpoint of a (Pragmatic)
Behaviorist**

Lloyd G. Humphreys

University of Illinois, Urbana-Champaign

Intelligence from the standpoint of even a pragmatic behaviorist may seem like a contradiction in terms for many readers, but “proof of the pudding is in the eating.” Any behavioristic approach seems outmoded in today’s climate of opinion in which cognitive psychology is ascendant, but psychology is still the science of behavior, not of cognition per se. Our task is to explain and predict behavior, with or without hypothetical mental constructs.

Behaviorism was emanant in the building of laboratories by universities. The movement was well underway at the University of Chicago in the writings and research of Dewey, Angell, and Carr prior to the arrival of Watson.

The writing of this manuscript was supported by the Department of Psychology, University of Illinois, Urbana-Champaign. A nearly identical manuscript is being published with commentary in *Psychological Inquiry*. Lawrence A. Pervin, Editor. This arrangement has been approved by both editors.

Following Watson, behaviorism developed many diverse shoots. The Gestalt movement quickly became incorporated in a generic American behaviorism. The common striving for objectivity, quantification, and control was an important legacy. I have no tolerance for the point of view that rejects for psychology the model of research of the physical and biological sciences.

A pragmatic behaviorist could also acquire serious reservations about a great deal of psychological theory. Too many theories do not lead to specific predictions of research outcomes that can potentially be disconfirmed. Global theories are especially suspect. Theories should start with good data and are necessarily small in scope in the beginning. Many psychological theories try to explain too much and are unfortunate heritages of our philosophical past. William of Occam's razor is directly applicable to psychological theory: Do not develop unobservable constructs unless they are required to explain the phenomena, and do not use constructs that cannot enter the development of testable hypotheses.

There is also a problem of the vocabulary to be used in psychological theories. Our vocabulary is shared with the humanities, theology, law, and people at large to a greater extent than other sciences. There are two realistic choices: either adopt a new vocabulary, or redefine words as necessary for scientific purposes. A third alternative that requires molding psychological theory to fit the definitions of the man in the street is not acceptable.

THE BASIC APPROACH

Intelligence will be defined as a phenotypic behavioral trait. A phenotypic trait is an observable characteristic, such as height or length, of a biological organism. Some traits can be directly observed and evaluated subjectively, others require instrumentation. Physical traits differ in most cases from behavioral ones, however, in being easier to define and measure. The former tend to stand out perceptually from each other so that agreement on what is being measured is more readily attained, but there are exceptions. The distinction between systolic and diastolic blood pressure and the importance of measuring each required both research and conceptual development.

Measurement Methodology for a Behavioral Trait

Once a phenotypic trait has been tentatively defined and a measure proposed, certain functional properties of the measure of the trait must be investigated. Among these are the reliability of the measurements, the robustness of the measurements under a variety of conditions, and the stability of measurements over time as maturation and learning take place. Because behavioral traits are

generally measured by psychological tests, a fourth category, the homogeneity of the test items, must be added. Each item must measure a component of variance common to all of the items that have been proposed as indicators of the trait. The trait variance in the items, as long as it is nonzero, is not required to be of any minimum size. If it is small, the trait variance in the total score on the test can be pushed to an acceptable level by adding the needed number of parallel items. High correlations among items are not required, but high homogeneity of the total score (coefficient alpha) is needed for satisfactory measurement of a hypothetical trait.

Given a research basis for measurement of the trait, the next step is to study the correlates of the trait. Correlations are obtained with other traits, including both physical and behavioral ones, and with experimental treatments. Without knowledge of measurement characteristics, interpretation of these correlations is uncertain. Also, if it is possible to control environmental variation adequately, the genetic contribution to the total variance of the trait can be obtained.

If the estimate of heritability is greater than zero, covariances involving the trait may also have a genetic component, and their interpretation affected. Even if the controls are good enough to justify a dependable point estimate of heritability, the estimation of an individual's genotype is made with a great deal of error for any heritability coefficient appreciably less than unity and adds nothing to the information about the individual furnished by the observed phenotypic score.

The goal of defining intelligence as a phenotypic behavioral trait is not a substitute for understanding *how* people or other biological organisms solve problems that are said to require intelligence. Neither is my goal a substitute for understanding the anatomical and biochemical mechanisms underlying the behavioral trait. These problems are legitimate and important, but they supplement the present approach. Research in those areas, if it is to be related to a behavioral trait, requires definition and measurement of the trait.

Traits and Causality

A word about attributing causality to a present trait for some future achievement is in order. Height is positively correlated with achievement in basketball among adults, but there are many exceptions. No one expects height to be associated with all of the variance in basketball achievement. Predictions are probabilistic, and probabilities change as measures of other traits are added. What is aptitude for playing basketball? Is it height, or is it the most predictive combination of traits that can be assembled? Whether the first or the second of the preceding two definitions is used, the correlation with adult achievement at 18 is a good deal smaller when the trait is measured at age 6 than at age 18. Furthermore, it is not merely the training that intervenes between 6 and 18

(environmental variance) that reduces the correlations. Relative standing within an age group is not a stable characteristic of a person for most physical traits. Perhaps aptitude serves to obfuscate rather than clarify thinking about traits, and should be dropped from our vocabulary.

Outline of the Discussion that Follows

I shall define intelligence, describe how the construct is organized and measured, make some assumptions about its bases in development, and then describe numerous hypotheses inferred from the ground work thus laid. I shall conclude with remarks concerning the importance of the construct in human activities.

THE CONSTRUCT OF INTELLIGENCE

I shall start with the definition of intelligence that appeared some years ago (Humphreys, 1971). Intelligence is the acquired repertoire of information, knowledge, and intellectual (cognitive) skills available to a person at a particular point in time. Individual differences in intelligence are monotonically related to the size of this repertoire. To avoid circularity, intellectual is defined by the consensus among experts working in the area. The repertoire is acquired during development, but it is acquired by a biological organism. Thus there is both a genetic and an environmental substrate for the trait.

Measuring the Repertoire

The intellectual repertoire can only be known behaviorally: The examinee must be both able and willing to produce the behaviors needed. Thus motivation of examinees to do well is critical. Physical characteristics such as blindness and deafness and psychoeducational characteristics such as bilingualism complicate the problem of assessment.

The repertoire can be assessed in two different ways: with ratings made by persons who have had ample opportunity to observe the relevant behavior of the examinee, and by tests. Ratings made by different raters and scores obtained on different tests are positively intercorrelated for each method and are also correlated with each other. There are also unique components in ratings from rater to rater, in scores from test to test, and from rater to test. These unique components tend to be larger in ratings than in tests because sources of error are easier to control in tests.

The variability in ratings has many sources, but includes the kinds of opportunities to observe that raters have had and the implicit definitions of intelli-

gence adopted by the raters. Ratings of classroom teachers tend to be more like each other than they are like those of athletic coaches or playground supervisors. Uniqueness can also be reduced by instructions to raters to exclude from their definition of intelligence traits of character, work habits, sensory acuity, and motor coordination. Specifying a standard test of intelligence such as the Stanford-Binet or a Wechsler scale also reduces uniqueness on the test side of the equation. As data are obtained under more controlled circumstances, correlations between ratings and test scores converge on the same construct.

Inequality of Opportunity

The measurement of intelligence has been complicated by loose thinking concerned with opportunity to acquire the behaviors in the repertoire. Some have defended tests of intelligence on grounds that opportunity of potential examinees in a particular culture was approximately equal. Others have rejected the tests because opportunity was clearly unequal. Still others have tried to devise items that would be culture-free. All have missed an essential point: It is the phenotype that is being measured. When height is measured and used for any purpose, no assumption is necessary that there were no nutritional differences among persons during development. Performance in life is a function of phenotypic traits, not estimated genotypes. This may not be fair, but the lack does not invalidate measurement or theory. The behaviors observed and measured when assessing the phenotypic trait of intelligence are acquired by a biological organism dependent on the luck of the draw genetically. Development also occurs in widely different environments and in environments that are constantly changing. Opportunity is obviously not equal between or within any grouping of a population or for a given person over time. Whether phenotypic traits can be modified by environmental manipulation, by how much, by what methods, in what period of time, at what point in development, and at what costs are appropriate research questions.

Exposure to Content

Of course there is an experiential requirement for a valid test, although equality of experience is not necessary. If the test is in English, familiarity with that language is required. If an examinee is multilingual, he or she should be asked each question in the two or more languages and be given credit for a correct answer in any one. An intellectual repertoire can be acquired in two or more languages and the examinee can be unable to express the element in any language except that in which it was acquired. Beyond this, the examinee must have been exposed to the information, knowledge, and skills. Exposure requires little more than presence in the environment. Few dropouts occur prior

to age 14. After spending 8 years in a nation's school, the requirements of exposure to the contents of intelligence tests have clearly been more than met.

ORGANIZATION OF THE REPERTOIRE

Elements of the repertoire form a positive manifold (intercorrelations are positive), which is the sole basis needed for the use of "general" to describe intelligence (Humphreys, 1979). The correlations among individual acts or items are, however, quite small. Intelligence is general, but at this level of detail the amount of specificity in any element is far and away more in evidence than is generality. When test items are linearly combined to form a total score, or when individual acts are summated in the form of ratings, grades in school, performance in industrial or military training, or on-the-job proficiency measures, correlations among the aggregates of the elements become much higher. There is still unevenness in the level of performance of persons from one aggregate measure to another, but generality is now substantial. For example, the correlation between total score on verbal and quantitative tests is larger than the item correlations within either of the tests. Even so, verbal items are more like each other than they are like quantitative items.

Factors in the Repertoire

The considerable variation in the size of the correlations among aggregates of the repertoire is systematic so that multiple factors can be defined. The positive manifold supports the use of the modifier "general" to describe intelligence; the systematic variation in size defines multiple factors. That the mathematics of factor extraction defines orthogonal factors from this systematic variation has been widely misinterpreted. There have also been widely disparate claims concerning their number. Unfortunately, the number depends on the selection of the tests to be factored. I have described (Humphreys, 1981) a compelling case for the numbers to be in the thousands if one starts with a wide enough sample of tests and a large enough sample of examinees so that small differences in correlations become stable. This analysis indicates that a continuing search for mental elements by means of factor analysis represents a dead end psychologically.

There is a second way of describing the systematic variation in size of the correlations in the intellectual repertoire that is independent of ordinary factor analysis. This is the radex model of Guttman (1954). To a first approximation, the elements of the repertoire can be located in two-dimensional space in accordance with a small number of assumptions and principles. Each aggregate measure is assumed to be error-free, and the sample of measures is assumed to

be intensive as well as extensive. Intensive sampling has the effect of providing little room for nonerror specific variance, unless specificity is characteristic of the measure as distinct from being a function of the sampling of tests. The cognitively most complex measures are grouped closely together in the center of the space whether their content is verbal, numerical, or figural. Variation in complexity is portrayed along radii leading out from the center with the simplest measures in the periphery. True specifics are due to superficial aspects of measures and are, therefore, cognitively simple. Content does have an effect on the size of correlations so that verbal, quantitative, and figural items are clustered in pie-shaped segments of the area of the space. Distance between points is inversely related to the size of the correlations between measures. For example, a measure of quantitative reasoning would be close to the center of the space, and accuracy in clerical number checking would be in the same segment of the space but close to the periphery.

Defining a General Factor

The intercorrelations of the elements of the repertoire define a general factor on which every element has a positive loading. Also, when the contribution of the general factor to covariation is held constant, there is no residual negative covariation. The size of the loadings varies with the variation in size of the mean correlations of the individual elements, which in turn are a function of complexity. Given adequate methodology that I have described elsewhere (Humphreys, 1982), the general factor can be defined uniquely from one sample of intellectual measures to another and from one sample of the general population to another. (Restriction of range of talent has effects on factors that are beyond the scope of this discussion.) The general factor represents a substantial amount, but not all, of the common variance in the repertoire.

The construct of general intelligence does not necessarily represent anything more than a mathematical dimension. As such it is "real" even though it cannot be observed under a microscope. Knowing that there is a highly replicable general factor that describes a substantial proportion of the covariance among the elements of the repertoire is both a necessary and a sufficient condition for the construct and for its measurement.

Tests of the General Factor

The general factor of intelligence cannot be measured directly because each and every item contains variance that is orthogonal to the general factor. The latter can only be estimated by a test, but given a sufficient number of items that sample the domain with sufficient breadth, the "noise" in the items contributes progressively less to their linear combination (the total score) as the

number of items increases. It may seem paradoxical, but the variance of the total amount of noise shrinks, and the variance associated with the central construct increases as the number of items increases, as long as each item measures the general factor to an appreciable extent and as the heterogeneity of the residual noise is maximized. (For a fuller discussion see Humphreys, 1985.) The indeterminacy in estimating the general factor can approach zero closely if there is a sufficient number of carefully selected items.

The two-dimensional plot of the elements in the repertoire serves as a guide to the selection of the most effective types of items to include in an intelligence test. Items close to the centroid of the space have the highest loadings on the general factor so that a relatively small number of them provide a valid estimate of that factor. Tests of knowledge of the meaning of abstract words, arithmetic reasoning, and spatial reasoning are close to the centroid. Items like these appeared in the original Binet-Simon scale.

Alternative Samples from the Repertoire

Although the types of items included in so-called aptitude and achievement tests are not identical, many tests labeled "achievement" lie very close to the centroid of the space. Reading comprehension is the example par excellence. If a composite is formed from the individual tests of a standard achievement series, the correlation of that composite with a standard test of intelligence will be almost as high as the correlation between two standard tests. An achievement composite not fully representative of the space obtained from the 12th-grade data for Project Talent has a correlation with the intelligence composite of .850 and .825 in males and females, respectively. These correlations are evaluated in the light of the reliability of the intelligence composite, which is estimated as .899 and .887 in the same two samples.

If one selects tests from an array somewhat removed from the centroid and obtains a composite, another high correlation is obtained. Nonacademic information tests are not near the centroid, but a composite of such tests from the 12th grade of Project Talent, also not fully representative of the space, has correlations with the intelligence composite of .787 and .777 in the two sexes, and the correlations with academic achievement are .826 and .814. Furthermore, most of the nonacademic information tests show substantial sex differences in both directions, although there are more tests showing male than female superiority. Virtually every item contains a substantial systematic bias component of variance.

Within each sex, as well as between sexes, there are wide differences in the opportunity for a given person to acquire information about such topics as the Bible, hunting, fishing, dance, exotic foods, and so on. Equality of experience with the content of the items is not required in estimating the size of the reper-

toire if the sampling is broad enough. Although it has not been done, a composite formed from an even larger number of tests widely scattered around the periphery of the space and heterogeneous in methodology as well as content would produce similar correlations. Such a test would not be practical either in time or conditions to administer, but it should be formed experimentally. Because tests of short-term memory are prominent in the periphery, obtaining the predicted outcome would have theoretical significance.

Ratings in the Test Space

Ratings of intelligence are also in the space defined by the test scores. Their location depends on several factors. If the rater does not know the ratee well, or if the acquaintance is based on a limited behavioral sample, the ratings are likely to be peripheral. If raters are allowed to select their own individual definitions of intelligence, but are otherwise well qualified to rate, ratings will scatter in the space as a function of the definition adopted. If well-qualified raters are instructed to disregard desirable qualities of the ratee such as ambition, hard work, emotional adjustment, and morality, that are not a part of the intellectual repertoire, and rate on use of symbols, solving problems, and so on, ratings will be closer to the centroid of the space. Each of the ratings made *independently* by n qualified raters will have a substantial correlation with the test of intelligence and the aggregate of the n ratings will be more highly correlated with the scores on the standard test than the median of the individual ratings. The gain will depend on the independence of the raters and the breadth of the behavior sampled.

Performance Measures in the Test Space

Academic grades that are based on proficiency in the subject matter lie fairly close to the centroid of the test space, but are also affected by content. English grades lie in the verbal segment, arithmetic grades in the quantitative segment, and engineering grades in design in the visual-spatial segment. If letter grades had the same meaning for degree of proficiency in the course for all courses and all instructors, which they decidedly do not, a composite of grades would be closer to the centroid than grades in any single course. This does not deny that grades based on proficiency include other variance, such as motivation and hard work, but nonintellectual traits have limited impact in the short run on level of proficiency attained in a new cognitive task. Individual differences on the task are not static over time, but change in the cognitive domain takes place slowly.

Measures of proficiency in occupations, including supervisor and peer ratings, that include tasks from modest to high levels of complexity also appear

in the test space. If performance on the job is cognitively complex, one knows in advance that correlations of that measure with scores on an intelligence test will be positive and of a size influenced primarily by the reliability and validity of the performance measure. Performance in different occupations such as mechanical maintenance and repair, on the one hand, and clerical occupations, on the other, fall in different segments of the test space but are close enough to the centroid in their respective segments that correlations with general intelligence are quite substantial.

Standard Tests of Intelligence

Although there are many tests that are labeled intelligence, two stand out from the rest in terms of their item sampling as good approximations of the general factor. These are the Stanford-Binet and the various Wechsler scales. Based on content they can be called standard tests of intelligence, but both share a limitation. Neither furnishes a useful score that measures the intelligence herein defined. Intelligence is the size of the repertoire, or the level on the general factor that reflects the size of the repertoire. Both of these tests were designed to furnish IQs, which are measures of relative intelligence and do not reflect growth.

When IQs were computed by dividing mental age by chronological age, mental age was a measure of intelligence. An IQ is a useful addition, just as knowledge of height relative to an age group is a useful addition to knowledge of height. Mental age units were abandoned because it was difficult to meet the requirements of an invariant scale for IQ with their use. The lack of a replacement is a serious gap because it confuses ability to perform with performance relative to age. For children, one recourse is to multiply the obtained IQ by the child's chronological age. Mental age, of course, was also not useful for adults, but a common scale for all age groups that would reflect the size of the repertoire could be formed. It is fortunate that test publishers retained grade-equivalent scores for achievement tests in spite of psychometric criticisms of their measurement properties. They do define a scale on which growth can be measured.

The standard tests and their close relatives obtain high marks on the four criteria described under the earlier heading of measurement methodology. There is no problem in measuring estimates of the general factor reliably as long as sufficient time and items are devoted to the task. Measuring a behavioral trait reliably is time-consuming and requires standardization of critical conditions, but such measurement does not suffer otherwise in comparison to the use of physical scales. If examinees are well motivated, a critical condition, measurement is also remarkably robust to conditions in the examinee and the environment that are considered stressful. Among these are minor illness,

fatigue, loss of sleep, ambient temperature, and ambient noise. These are not excuses for measuring intelligence unpleasantly, but they do show that the trait itself is robust.

There is no requirement that a behavioral trait be fixed and unchanging just as there is no similar requirement for a trait of physique. Information about stability is required for theoretical and applied interpretations and is available for standard tests. Finally, the standard tests are approximately homogeneous with respect to general factor content, but only because they are so heterogeneous with respect to the content introduced by their minor factors. The traditional criticism that a given score can be attained by different patterns of response is beside the point as long as the score is used and interpreted only as an estimate of the general factor.

It is fortunate in many ways that the two standard tests do a reasonably good job in estimating the intellectual repertoire. Item types have been sufficiently constant over 80 years so that the large number of correlates of scores of these tests and their close relatives can be considered comparable over the same time period. At least for the foreseeable future, it is more fruitful to retain the standard tests and to redefine what developers and most users thought the tests were measuring than to continue the search for the Holy Grail, the "real" intelligence.

ASSUMPTIONS ABOUT DEVELOPMENT

My basic definition of intelligence required a single paragraph. The subsequent discussion depended very largely on empirical observations of correlations among behaviors in fleshing out the construct. Still needed, however, are some assumptions about human development.

The Genetic Substrate

The genetic substrate for general intelligence is polygenic. In a limited sense, because intelligence is a behavioral trait, the substrate includes the genetics of the whole organism. Obviously, this point of view can be pushed too far. On the other hand, it is simplistic to assume that the genetic substrate is restricted to anything like a bodily organ. It seems more reasonable that the genetic substrate is primarily that for the central nervous system, but the entire sequence from stimulus through to response should not be forgotten. The assumption that many genes are involved must be taken very literally.

An assumption that the genes in the substrate for intelligence do not all "fire" at the moment of conception is also reasonable in the light of numerous

late-appearing genetic effects. Relative individual differences in physical traits are not stable during development and decay.

Although a great deal of evidence concerning the heritability of intelligence has been published, behavioral geneticists have not considered adequately whether group factors in intelligence have a specific genetic substrate or whether the genetic contribution was restricted to the general factor content of the measures of group factors. General intelligence tests cover the broadest spectrum of behaviors, and every group test of a narrower factor such as verbal comprehension necessarily includes general factor variance.

Cardon et al. (1992) did claim substantial heritabilities for verbal, spatial, and memory group factors, but their general factor was estimated by a methodology that underestimated the variance of the general factor. Humphreys (1971), using a different methodology, obtained evidence for the heritability of a spatial factor only in his sample of male twins.

Still narrower factors defined by the difference between a measure's communality and its reliability (specifics) can be confidently assumed to be environmental in origin. Finally, item variance in standard tests of general intelligence is largely the product of multiple, independent environmental determinants.

The Environmental Substrate

Elements in the repertoire are acquired. Learning is accordingly implicated, but environmental influences can also have indirect effects. Both types of influences are multiple. Polygenic determination is accompanied by polyenvironmental determination. These separate effects will each be discussed briefly.

A detailed theory of learning and motivation is not required. If learning takes place effectively, elements will be added to the repertoire more rapidly than when the process is ineffective. If the content to be learned is cognitively complex, general intelligence will benefit more than if the content is simple. This does not mean that all persons acquire new elements in the repertoire with equivalent speed and sureness. As described earlier, there is a genetic substrate for the acquisition of the intellectual repertoire.

There is also an organic substrate that is environmental in origin. Differential experience produces differences in dendritic growth. It also seems probable that early experience is more effective than experience that comes later in maturity. The extended period of infancy and dependency in the human is accompanied by a high degree of plasticity but this property wanes with increasing age. In addition, there are the accidents, diseases, and other physiological pathologies that affect the biological organism and may well affect in turn the acquisition of the intellectual repertoire (Lubinski & Humphreys, 1990a). The problem of testing the repertoire in the physically handicapped

was mentioned earlier, but the handicap can affect acquisition as well. The objective in testing physically handicapped persons is to sample their repertoires, not to estimate their genotypes. Perinatal anoxia and early nutrition may present similar problems.

Size of the Repertoire

By the age of two, the repertoire is already large. Responses to verbal stimuli preceded use of words, and nonverbal problem solving has been prominent for many months. Being attentive to changes in stimulation revealed by turning of the head seems to belong in the intellectual repertoire. A student and I, as a matter of fact, published data congruent with the hypothesis that the general factor can be measured at 12 months (Humphreys & Davey, 1988). Although individual differences on measures of the factor change fairly rapidly in the early months and years, the change appears to be gradual and lawful. From this point of view, the general factor gradually evolves into its later manifestations.

Growth is a process of continuous accumulation. With almost 100% of American children in school from age 6 to 14 and exposed to a broad, relatively homogeneous curriculum, growth tends to take place on a broad front. Public education also requires continuous practice of earlier elements. Differential experience in the home and differential reinforcement of early school experiences provide a possible basis for the initial differentiation of group factors as well.

The high school curriculum provides somewhat greater opportunity for specialized growth, although this is still around a common core. Additional group factors do become defined during this period (Atkin et al., 1977). Specialization becomes dominant in postsecondary education and growth in general intelligence slows markedly. By age 18 the intellectual repertoire is huge.

Relation Between Repertoire and Gain

There is ample basis in the discussion of the genetic and environmental substrates for intelligence for the following conclusion: True score gains between times 1 and 2 are necessarily imperfectly correlated with true score of the repertoire at time 1. Another way of stating this involves the distinction between reliability and stability over time. The latter statistic will always be smaller than the root of the product of the reliabilities at times 1 and 2.

The correlations of scores at base periods with gains and the correlations between successive gains can be quite small. Genes that "fire" late in development are determiners of bodily structures and functions, but being independent of other determinants in their action, they introduce seemingly stochastic changes in development. Similarly, the environment is constantly changing and

to a large extent in a stochastic manner insofar as the individual child is concerned. Many years ago, following a proposal by Anderson (1939), Roff (1941) computed raw score gains with raw score bases and largely found small negative correlations. Because there is a bias in raw score data toward negative values, it seems probable that the true score correlations might have been no larger than small positives.

IMPLICATIONS OF THE THEORY

At this point it is appropriate to discuss a theory about intelligence as opposed to a definition. As such it should lead to testable hypotheses about relations involving the trait. These relations constitute only a limited segment of the phenomena subsumed under the rubric of intelligence by psychologists. The process by which elements of the repertoire are acquired or how the repertoire is used in solving problems are clearly important, but outside the scope of the theory of the trait.

Data are available for many of the testable hypotheses about the trait. Eighty years of testing children with tests that differ little from what I have called the standard ones have provided a wealth of information. In an important sense the theory was derived from the large body of data that has been accumulated. This approach may differ from a great deal of psychological theorizing, but it provides a coherent structure for the data and leads to additional research. It also reflects the biases concerning psychological theory that were discussed earlier.

Hypotheses that follow from the theory will be discussed under three headings: group means and changes in group means, stability and instability of relative intelligence of individuals, and concurrent and predictive correlates of intelligence. Space will not permit a full discussion of the evidence relevant to each hypothesis. The hypotheses that follow depend to some degree, heavily in many cases, on the absolute size of the typical repertoire as it changes during development.

Group Means

These hypotheses are admittedly difficult to test with confidence. Imposition of adequate experimental controls, of which random assignment to treatments is crucial, is not feasible. Statistical control is never completely adequate. Ability to test latent trait models was an important advance, but such models only control the attenuation introduced by measurement error. There remain more important problems of the construct validities of the latent traits identified and whether an important one has not been identified.

There are also problems with respect to units of measurement. Has Item Response Theory (IRT), as implemented by several different computer programs, solved the problem of equal units for the measurement of growth of intelligence or even a narrow element of the intellectual repertoire? I remain skeptical. Mental age and grade equivalent metrics are useful, but the units of measurement are widely known to be unequal. Measures of relative intelligence within a group homogeneous with respect to chronological age have only limited uses.

Nevertheless, the problems of general intelligence are important. Inability to do controlled research should not be the sole determiner of where resources should be expended. Careful investigators need to do the best research possible from the sampling of defined populations and choice of measurement instruments to sophisticated use of modeling. In the end, furthermore, conclusions should be as tentative as the data require.

1. *The mean level of intelligence in a population can change over time.* A substantial increase in the population of young adults tested in World War I and World War II was reported by Tuddenham (1948). More recently increases in the perceptual-visual-spatial component of the repertoire were reported by Flynn (1987). Critics use such data to denigrate intelligence tests, claiming that "real" intelligence did not change, but their construct cannot be inferred from any measurement operation. Such critics are implicitly hereditarians.

2. *When the size of the intellectual repertoire changes, there will be corresponding changes in educational achievement and occupational proficiency.* World War II draftees would have been more able than their World War I counterparts to acquire the cognitive knowledge and skills required in their military assignments. To test this hypotheses it is essential to distinguish between a gain in score based on practice or coaching on a given set of items and a gain in the total repertoire. Current discussions concerning the importance of functional literacy training for employment and industrial productivity assume the accuracy of the hypothesis.

3. *Psychologically important gains in intelligence will occur only with the expenditure of substantial effort in time and resources.* If the gain between the two world wars was due to the increase in years of education, as it probably was, consider the economic cost of keeping a growing population of children in school for approximately 3 additional years. One cannot expect to obtain a large gain in intelligence from an experimental program whose duration is measured in months.

4. *For a given level of effort there will be greater effects on young children than on older children.* Growth in intelligence is measured by growth in the intellectual repertoire. Producing an increment to that repertoire by a given intervention can be accomplished more easily when the repertoire is relatively small.

5. *For effective growth in intelligence there must be a continuous supportive psychosocial substrate.* Preschool intervention programs typically have short-term effects. There must be continuing exposure and continuing effort, but the latter requires effort on the part of the learner as well as by the society. Social effort that does not affect individual effort is not sufficient.

6. *Changes in intelligence are a function of the kind of intervening educational experiences.* In the system of education studied by Hamquist (1968) students could be grouped in the following categories: compulsory level, vocational, lower secondary, and gymnasium. Estimated gains, after allowing for either imperfect stability over time or imperfect reliability on pretest and posttest, increase monotonically from compulsory level to gymnasium.

7. *The chronological age at which the growth of intelligence in the general population essentially levels off depends on the educational practices of the society.* There is no age set by biological growth processes. Terman used 16 as the maximum age in computing IQs in 1916, 18 in 1937. Today, however, we no longer try to measure intellectual growth, so today's "maximum" is unknown.

8. *In the absence of specific biological deterioration, there will be little loss in intelligence with increasing age.* There is little forgetting of overlearned and continuously practiced skills. Some elements may enter and later disappear, but the total will rarely decrease. This hypothesis does require the distinction between intelligence and relative intelligence, a distinction that is frequently neglected. Thus, monotonic decrease in the mean IQs of rural southern black children (Jensen, 1977) does not negate this hypothesis.

9. *Intervention can change the level of an individual component of intelligence more rapidly than the total.* Even so, some components such as aural and visual comprehension of language, which are themselves highly correlated, are so central to the repertoire that obtaining more growth than typically occurs is slow at best. The intervention problem is quite different, however, among the small number of individuals who are substantially retarded in reading as compared to their level of aural comprehension.

10. *The components of general intelligence do not all grow at the same rate.* The age of leveling off varies from one component to another (Horn, 1989). These differences are correlated with differential exposure to the different types of content. Engineers and physical scientists, on the one hand, and humanists on the other, will show different patterns of growth and decline for verbal, quantitative, and spatial-visual components. The last of these is prominently involved in measures of so-called fluid intelligence.

11. *There are mean differences in intelligence among groups defined demographically.* These groups do not experience identical environments, and there may be genetic differences as well. Groups defined by level of socioeconomic status have a genetic component to the variance of their scores on standard tests of intelligence. Groups whose gene pools have been partially segregated

historically are likely to be genetically somewhat different. Whatever the causes may be, group differences have social consequences.

12. *Among adult representatives of groups demographically defined it is difficult to overcome existing deficits in general intelligence.* This hypothesis is independent of how one views the nature–nurture problem. It is not easy to compensate in the adult for low levels on a phenotypic trait, even if the variance of the trait is largely or entirely environmental in origin. Observe the many problems encountered by adults in the acquisition of a second language. The popular view of intelligence is one of innate power or capacity that can be released at any time by providing opportunity and that quickly comes to full flower. There are no data to support this view. The intelligence I have defined develops slowly with support from the society and with effort on the part of the learner.

13. *For effective growth in intelligence, the specific curriculum and standards of achievement should be set somewhat above the current mean level of the group.* Individuals in the group must be currently motivated or able to be motivated to work at the learning of the material, whatever it may be. If the students within the group are highly heterogeneous in preparation for learning the material, both the highly prepared and the poorly prepared are disadvantaged. It is especially disadvantageous to be so far below the mean that competition is essentially impossible.

Stability of Individual Differences

Hypotheses in this area are concerned with change in relative intelligence over time. Time is not, of course, the effective variable, but in the absence of information about the actual determiners, time is the appropriate dependent variable. Longitudinal research over time does require examinees who are relatively homogeneous in chronological age. Thus it is immaterial whether a growth or relative score is used.

1. *Scores on a test of general intelligence intercorrelated over occasions will produce a quasisimplex matrix.* Such a matrix tends to have its smallest correlation in the upper right-hand (or lower-left) corner, and the largest correlations adjacent to the principal diagonal. If the obtained correlations are corrected for the attenuation produced by measurement error, the preceding statements can be made without the qualification introduced by “tends.” If standardized tests are used, reliabilities tend to be relatively constant from occasion to occasion so that the quasisimplex pattern is obtained in large samples without reversals. Stability coefficients drop monotonically as time between occasions increases.

2. *After correction for attenuation the correlations will approximate a true simplex pattern.* The assumption of a zero correlation of true scores between

base at time i and gain between i and j may not be veridical but it has resulted in acceptable fits of observations to model (Humphreys & Davey, 1988; Humphreys & Parsons, 1979). Correlations with true score gains clearly do not approach unity, which is required for a fixed intelligence.

3. *Stability over time is a function of the initial age of the examinee.* As the repertoire increases in size, it becomes increasingly difficult to obtain a gain proportionate to the size of the base. The corollary of this hypothesis is that a gain of the same absolute size will have a larger effect when the repertoire is relatively small. It also follows, in the absence of special interventions, that the stability of intelligence over a given unit of time should increase throughout maturation.

4. *Low correlations between intelligence in the first few years and adult intelligence reflect the normal growth of intelligence.* This is contrary to the common assumption that early tests do not measure intelligence, but this alternative is viable (Humphreys & Davey, 1988).

5. *Instability over time is as characteristic of physical traits as of intelligence.* While there is a psychosocial substrate for height and weight, the genetic substrate for both, especially height, is more important than that for intelligence. The stability of height from year to year during development may not be much greater, however, than for intelligence when correlations have been corrected for attenuation (Humphreys, Davey, & Park, 1985).

6. *Change is more rapid for the narrow components of intelligence than for the total score.* There may be differences from one component to another, but the results of simplex-fitting attempts by Humphreys, Park, and Parsons (1979) for 16 tests in male and female samples seem to be in line with this prediction.

7. *Change is a function of the intervening psychosocial substrate.* There should be more change for students in an academic curriculum than in a trades curriculum. Harnquist's data (1968) support the hypothesis, but uncertainty is introduced by differences in variance from group to group.

8. *A highly competitive learning environment will produce more change than a more placid one.* If students are able to deal with the competition, more gain overall will produce lower stability. Verbal and quantitative college board scores can serve as partial surrogates of general intelligence. Stabilities of these scores, if the problem of restriction of range can be solved, should be smaller at highly selective institutions than at state colleges (or universities) that only a few years ago were teacher-training institutions.

9. *Each continuously distributed psychological trait exhibits a modest degree of similarity among relatives.* Such traits have multiple determinants. Without regard to the genetic-environmental mix, differences among relatives produce resemblance coefficients that are a good deal less than unity after correcting for attenuation. Lack of identity requires that children of extreme parents regress toward the mean on each and every such trait as do the parents of extreme children. For general intelligence, even though there is a good deal of

variability reported, a correlation of .50 between either of the two parents and a single child is a reasonable estimate. More distant relatives are more distant both genetically and environmentally, so resemblance coefficients become smaller with increasing distance.

Validities of Intelligence Tests

Test validities are usually described by correlations as are the stabilities of individual differences in intelligence. Thus there is a good deal of similarity among hypotheses in the two areas.

1. *Proficiency of any performance in education, industry, and the military requiring intellectual (cognitive) tasks is correlated with scores on a standard test of intelligence.* Intellectual is again defined by consensus, and proficiency is represented by an aggregate measure, not a single act. The "validity generalization" doctrine was established empirically (Hunter, 1980; Schmidt & Hunter, 1977), but is readily derived from present theory. Educational and occupational tasks that broadly sample the repertoire are necessarily highly correlated with a test that also broadly samples the repertoire.

2. *Maximum predictive validity is obtained when the repertoire sampled by the test matches the repertoire sampled by the criterion performance.* Performance in a blue-collar mechanical occupation is predicted more accurately by tests of mechanical reasoning and spatial-visual problem solving than by a test of general intelligence.

3. *An R-matrix of measures of proficiency obtained over multiple intervals of time will form a quasisimplex matrix.* Neither achievement nor so-called aptitude ever become stable.

4. *Performance measures at 18 are predicted with increasing error the earlier the age of administration of the predictor test.* As predictors of a child's intelligence, parental intelligence and socioeconomic status are only moderately valid (.50 and .40, respectively), and the child's intelligence at age 2 is no more accurate. If a program is widely adopted in which parents are expected to start saving for the expenses of higher education during the child's preschool period, problems will be created. Should all such children be admitted to college if they wish to attend and have a high school diploma? More importantly, should children whose parents made no attempt to save be discriminated against?

5. *Predictions from intelligence tests at a given point in time are unlikely to remain at a constant level of accuracy for early to late performance measures.* If neither the predictor nor the criterion is stable, correlations between the two are unlikely to be stable.

6. *A quasisimplex of criterion performance can arise from decreasing cognitive complexity of performance with practice.* The increasingly narrower sam-

pling of the cognitive repertoire produces late performance specific to itself.

7. *A quasisimplex matrix can also be explained by a gradual shift in the abilities required to perform well.* To test this explanation, a set of predictors is required differing enough from each other that increasing and decreasing gradients of predictive validities can potentially be found among the predictor tests. An increasing gradient is required for at least one test to support a change in ability.

8. *Decreased accuracy of prediction over time also requires information about the mean performance of the group.* Decreasing accuracy results in increasing amounts of regression of extreme persons toward the group mean. It is widely overlooked, however, that the group mean at the end of the time interval may itself be extreme with respect to the total population. The fundamental question is the following: Do the members of a selected group regress toward a superior mean or toward the mean of the population from which they were selected? The level of performance in a select group is hypothesized to remain high as long as there is appropriate institutional support. Thus the usefulness of the test whose predictive validity drops substantially with the passage of time cannot be evaluated solely by a correlation computed within the group. There are numerous essentially stochastic events that occur during a given interval of time, producing changes in individual performance, both up and down, that reduce the within-group correlation.

9. *Comparison of predictive and postdictive gradients of validities over time provides information concerning the kinds of changes taking place.* Humphreys and Taber (1973) found different shapes of these gradients for broad measures of abilities, on the one hand, and for the Advanced Test of the Graduate Record Examination on the other.

10. *Most predictive validities are gross underestimates of the importance of general intelligence in criterion performance.* As students move up the educational ladder from kindergarten to graduate and professional school, the successive restrictions of range of talent along the way are primarily on the general factor in intelligence. Although there are wide differences from institution to institution and from one curriculum to another within an institution, educational selection is substantial. For the high school class of 1962, for example, boys who completed the PhD were 1.26 standard score units above the mean of their 10th-grade class. Girls at the same level had a mean of 1.48. In the same high school class there is ample evidence for curriculum differences, but institutional differences were not available. Moving to the B. A. level in order to obtain adequate Ns for both sexes, boys are 1.28 units above the male mean in physics, and girls are 1.16 above the female mean in chemistry. In contrast, means in physical education are .06 and .43 for boys and girls, respectively. Approximate means in the several social sciences are intermediate, being .70 for boys, .95 for girls.

11. *Scores on an appropriate test have the same significance for criterion*

performance across different demographic groups. Appropriateness is judged in terms of the exposure criterion, not the equality of opportunity criterion. It is also judged in terms of the match in sampling the cognitive space of the test and the performance required. This hypothesis translates into an expectation of equality of intercepts in the regression of performance on the test after allowance is made for the effect on the intercepts of measurement error in the predictor. If an intercept difference remains, it is more productive to search for an additional valid predictor than for biased items in the present predictor.

12. *There are differences among socially important criterion performances in their predictability by an intelligence test or its most central components.* Among these are artistic, musical, and athletic abilities. The expectation, however, is not that predictive correlations will be zero, but they will be much smaller than those with achievement in the so-called learned professions. In some cases lower correlations are not produced by specific abilities in the periphery of the cognitive space but by traits more appropriately called personality or temperament. Success in certain kinds of selling or leadership are examples. The importance of specifying kind is indicated by the difference between leadership in a learned profession and in the command of an infantry platoon on the battlefield. Intelligence tests predict later ratings of officer effectiveness, but do so less accurately than they predict either pre- or post-commissioning grades in technical training. The effectiveness ratings are also known to be predictable by rated differences in personality traits.

IMPORTANCE OF GENERAL INTELLIGENCE

General intelligence is one of many human traits, but it is clearly an important one. It is highly related to educational success and to occupational attainment. The latter occurs in spite of inadequate educational guidance and inadequate support in our society for children of high intelligence from working-class and lower class family backgrounds. It is related to economic productivity, and although it is a difficult research problem, it is undoubtedly related to the ability of a democratic society to function as such. Remember, in this connection, that there is little difference between functional literacy and general intelligence.

Illustrating the Importance

The traditional method of presenting correlations between psychological tests, including those that measure general intelligence, and important social criteria has led the general populace, politicians, and reporters to underestimate the importance of intelligence. Individual differences correlations between intelli-

gence and various military performance criteria may, on average, be as high as .70, while correlations with academic grades in college, a restricted range of talent, may be represented by values around .40. In both cases standard errors of estimate are large so that there is a great deal of uncertainty in predicting individual performance.

Now consider an alternative. Divide the predictor test into multiple class intervals and compute a mean criterion score in each class interval. If there is adequate ceiling and floor for both the test and the criterion measure, the regression of criterion score on the test will be approximately linear. Now compute the product-moment correlation between the test and the criterion based on the means only, which is what engineers and experimental psychologists do when they wish to express the functional relation between two variables when each data point represents several observations.

When the preceding steps are followed, it is not unexpected in large samples to find a correlation of .99 for the means of tests and criteria. If there are 12 intervals, an r of .99 is greater than .94 at $p < .01$. The sampling error of a given mean is equal to the standard error of estimate in the distribution of individuals divided by the square root of the number of observations in the mean.

These correlations have interesting properties. They are independent of the size of correlations computed in samples of individuals, but they are dependent on the accuracy of the assumption of linearity. If the regression is linear, or approximately so, in the population, these correlations are also dependent on sample size. Given linearity and large sample size, the squared correlation in the sample of persons can be approximated quite accurately by the ratio of the variance of the criterion means to the variance of persons. However, the correlation based on individuals is not needed in making interpretations. The appropriate procedure is to make a judgment as to whether the gain in performance for each unit of change in the predictor in the raw score regression equation is important practically and theoretically. The question concerns the utility of the test in making decisions about the groups or in forecasting social trends. The utility judgment is not complicated by the need to balance the costs of false negative and false positive errors in prediction. The regression estimate of a performance mean is made with high accuracy in good data in both senses of the use of that term: an almost-zero standard error of estimate and a small sampling error.

Given high confidence in the gain in performance that can be obtained by a given increase in intelligence or a component ability, how does a democratic society deal with individual differences produced by the combined genetic and environmental determiners of intelligence? The problem may be in conflict with the egalitarian ideal, but not with the basic democratic value that stresses individual worth. Acceptance of the reality of the problem is the first step in dealing with it democratically. The next step is to know its manifestations and to deal with them rationally.

Instability Over Generations

As discussed earlier, parents and children are only similar, never identical, with respect to traits that are determined by many genes and many environmental influences. Given correlations in the population between a selected parent and a selected child of less than unity, and typically a good deal less, regression in both directions from one generation to the other takes place. If the correlation is .50, regression is halfway back to the population mean. If the generational differences involve relatives less closely related genetically and environmentally, the resemblance coefficient is substantially reduced.

Persons with an emotional bias toward equality in outcomes are due to be forever disappointed with respect to individual differences within and between present social strata, but need not be disappointed with respect to equality of outcomes of descendants two to three generations removed. Regression toward the population mean is progressive. The key is to maximize equality of opportunity in each generation. Genetics alone cannot produce a rigid class structure, but social inventions can.

Misconceptions about Heritability

In considering the social consequences of some degree of heritability of general intelligence, understanding the definition of the term is essential. Heritability is the proportion of genetic variance in the total variance of a trait in a defined population at a particular point in time. Heritability of intelligence in our species can vary from one population to another and from one time period to another.

As long as there is a genetic component in general intelligence, the goal in a democratic society is to reduce environmental variance as much as possible. Increasing opportunity for all citizens increases heritability. If the accurate figure were .80, as some have claimed, a democratic society could be proud of what it had achieved. If heritability in this country is lower today than it was a generation ago, as others have suggested, liberals have no cause for rejoicing. The probable explanation is that our class structure has become more rigid and that freedom of opportunity has decreased.

Mean levels do not enter the definition of heritability. Adding a constant amount to each person's intellectual repertoire by age 18 by strictly environmental means would not change the heritability coefficient. If we added a substantially larger amount to the intelligence of persons in the lower half of the distribution, heritability would increase.

Some conservatives have argued that a revolution in an existing aristocratic society would merely exchange one ruling class for another. Several comments are pertinent to this assertion. If the prerevolutionary society had a rigid class

structure, the new ruling class would be more able than the old. If the new rulers formed a rigid class structure, in the generation of their grandchildren or great-grandchildren the level of genetic endowment in the ruling class would approach the society's average. It is doubtful that environmental variance can adequately compensate. It also follows that highly able persons by this time have appeared in the lower classes. They are also frustrated by the shackles imposed on them by the class structure and are ready to lead a new revolution.

Perhaps the most dramatic evidence of genetic influence is typically overlooked. That it does not enter an estimate of heritability may be the explanation. The sheer number of children who exhibit high levels of talent and who are found in the most unlikely environmental circumstances is very impressive (Lubinski & Humphreys, 1990b).

Differences Among Demographic Groups

As I have defined intelligence, there is no doubt about the reality of demographic differences. A democracy must deal with these differences. Value judgments can and should override at times the implications of quantitative data, but such judgments should not ignore those data. Even worse is the strong tendency among many policy makers and policy shapers to substitute myths for data. More and better data are available concerning black and white comparisons, so I shall limit my discussion to those two groups.

An important empirical generalization is known and widely accepted among persons who read and can understand the research literature. When blacks and whites compete directly with each other and are evaluated by the same standard in education, industry, and the military services, blacks do not perform quite as well generally as the estimate from their supposedly biased test scores indicates (Hartigan & Wigdor, 1989; Linn, 1982). However, the differences tend to be about the size expected from measurement error in the predictor. That is, there is not sufficient evidence for a second deficit over and beyond that measured on a well-developed and well-researched predictive measure of ability.

It also follows that use of a lower standard for the selection of blacks in education, industry, and military service places blacks at a competitive disadvantage with whites on performance measures. I have described this practice as the strong form of affirmative action in comparison with the original definition. The latter required active recruiting of qualified minorities and selecting a minority person when qualifications were equal. It seems highly probable that strong forms of affirmative action are responsible for a substantial number of racial incidents and increases in racist feelings on the part of both blacks and whites (see Humphreys, 1990, 1991a, 1991b).

A correlation between distributions of means is also instructive for groups demographically defined. For example, two races, two ethnic groups, and three levels of socioeconomic status can be the basis for such a correlation with an N of 12. The 12 groups replace the class intervals of intelligence, and each group has means on both the predictor and the criterion. Correlations computed in this way will vary from test to test and from one situation to another, but they will be high, as long as the groups are exposed to the same content and are evaluated on the same standard. That is, mean test scores of races, ethnic groups, and social classes do predict quite accurately means of measures of educational and occupational performance. The plot of the linear regression of mean performance of mean test score will reveal errors of under- or overprediction, to the extent that such errors exist. Again, the regression equation in raw score form enables one to judge the importance of a given difference on the predictor.

There are, of course, many individual exceptions to group trends in all groups demographically defined. As a matter of fact, the distributions of individuals about both sets of means are approximately equivalent from group to group. These distributions about each mean in each set demonstrate the error of substituting a stereotype for accurate information about individuals and the need for the original definition of affirmative action. The high correlation between the two sets of means shows the consequences of the use of a strong form of affirmative action.

The important problem faced by a minority group is not a deficit on a predictor test, but a deficit in proficiency on a socially important performance. It is as difficult to identify black exceptions to the expectation based on test score as it is for whites. Discovery and use of tests that reduce the number of errors of prediction for blacks would be desirable, but such tests would do the same for whites. Present theory does not reject the possibility of an environmental solution to the problem, but it does strongly suggest that intervention should start early in pregnancy and be continuous thereafter. Entrance in the first grade is already late. Intervention should avoid the myths about self-concepts as prime determiners and focus instead on the acquisition of comprehension of oral and written language, writing, and skills in working with quantitative materials and concepts.

CONCLUSION

I have defined the phenotypic behavioral trait of intelligence that is measured by standard tests of intelligence. Genetic and environmental substrates affect the acquisition, storage, and retrieval of the intercorrelated intellectual (cognitive) knowledge, information, and skills that form a behavioral repertoire.

Intercorrelations define a factor common to the elements of the repertoire that support the designation of the trait as *general* intelligence.

With the addition to the definition of some assumptions about learning and development, a number of testable hypotheses can be derived. These are categorized into those that concern mean changes in intelligence in groups, changes in individual differences on the trait, and correlates with performances in various social roles.

Because intelligence is general, individual and group differences on the trait have widespread effects on the individual and on the society. General intelligence is far from being of sole importance in human affairs, but it is one of a relatively small number of important human traits. The number of traits that have appreciable effects on behavior outside the room in which tests are given is far smaller than the numbers that some ability and personality theorists have discussed.

Because intelligence is important, many liberals have felt pressure to reject any importance of the genetic substrate in the formation of individual differences in the trait. They have added an assumption, frequently implicitly, that environmental determination allows quick and easy compensation for earlier environmental deficits. The definition and theory discussed herein reject the latter assumption; and the theory does not require a resolution of the relative contributions of nature and nurture to the development of individual differences in intelligence. A democratic society can deal with the reality of an unknown, probably substantial, genetic contribution to variance. As a matter of fact, the goal should be to maximize that contribution by increasing environmental opportunities.

REFERENCES

- Anderson, J. E. (1939). The limitations of infant and preschool tests in the measurement of intelligence. *Journal of Psychology*, 3, 351-379.
- Atkin, R., Bray, R., Division, M., Herzberger, S., Humphreys, L., & Selzer, U. (1977). Ability factor differentiation, grades 5 through 11. *Applied Psychological Measurement*, 1, 65-76.
- 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.
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 271-291.
- Guttman, L. (1954). A new approach to factor analysis: The radex. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences*. Glencoe, IL: Free Press.
- Harnquist, K. (1968). Relative changes in intelligence from 13 to 18. *Scandinavian Journal of Psychology*, 9, 50-82.

- Hartigan, J., & Wigdor, S. (1989). *Fairness in employment testing: Validity generalization, minority issues, and the General Aptitude Test Battery*. Washington, DC: National Academy Press.
- Horn, J. (1989). Models of intelligence. In R. L. Linn (Ed.), *Intelligence: Measurement, theory and public policy*. Urbana: University of Illinois Press.
- Humphreys, L. G. (1971). Theory of intelligence. In R. Cancro (Ed.), *Intelligence: Genetic and environmental influences* (pp. 31–42). New York: Grune and Stratton.
- Humphreys, L. G. (1979). The construct of general intelligence. *Intelligence*, 3, 105–120.
- Humphreys, L. G. (1981). The primary mental ability. In M. P. Friedman, J. P. Das, & N. O'Connor (Eds.), *Intelligence and learning* (pp. 87–102). New York: Plenum.
- Humphreys, L. G. (1982). The hierarchical factor model and general intelligence. In N. Hirschberg & L. G. Humphreys (Eds.), *Multivariate applications in the social sciences* (pp. 223–240). Hillsdale, NJ: Erlbaum.
- Humphreys, L. G. (1985). General intelligence: An integration of factor, test, and simplex theory. In B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurement, and applications* (pp. 201–224). New York: Wiley.
- Humphreys, L. G. (1990). Review of a common destiny: Blacks and American society, edited by Gerald David Jaynes and Robin M. Williams, Jr. *Contemporary Psychology*, 35, 938–940.
- Humphreys, L. G. (1991a). Limited vision in the social sciences. *American Journal of Psychology*, 104, 333–353.
- Humphreys, L. G. (1991b). Some unconventional analyses of resemblance coefficients for male and female monozygotic and dizygotic twins. In W. M. Grove & Dante Cicchetti (Eds.) *Thinking clearly about psychology, Vol. 2: Personality and psychopathology* (pp. 158–187). Minneapolis: University of Minnesota Press.
- Humphreys, L. G., & Davey, T. C. (1988). Continuity in intellectual growth from 12 months to 9 years. *Intelligence*, 12, 183–197.
- Humphreys, L. G., Davey, T. C., & Park, R. K. (1985). Longitudinal correlation analysis of standing height and intelligence. *Child Development*, 56, 1465–1478.
- Humphreys, L. G., Park, R. K., & Parsons, C. K. (1979). Application of a simplex process model to six years of cognitive development in four demographic groups. *Applied Psychological Measurement*, 3, 51–64.
- Humphreys, L. G., & Parsons, C. K. (1979). A simplex process model for describing differences among cross-lagged correlations. *Psychological Bulletin*, 86, 325–334.
- Humphreys, L. G., & Taber, T. (1973). Postdiction study of the Graduate Record Examination and eight semesters of college grades. *Journal of Educational Measurement*, 10, 179–184.
- Hunter, J. E. (1980). *Test validation for 12,000 jobs: An application of synthetic validity and validity generalization to the General Aptitude Test Battery (GATB)*. Washington, DC: U. S. Employment Service.
- Jensen, A. R. (1977). Cumulative deficit in IQ of blacks in the rural South. *Developmental Psychology*, 13, 184–191.
- Linn, R. (1982). Ability testing: Individual differences, prediction, and differential pre-

- diction. In A. K. Wigdor & W. R. Garner (Eds.), *Ability testing: Uses, consequences, and controversies* (Part II). Washington, DC: National Academy of Sciences.
- Lubinski, D., & Humphreys, L. G. (1990a). *Some bodily and medical correlates of mathematical giftedness and commensurate levels of socioeconomic status*. Unpublished manuscript.
- Lubinski, D., & Humphreys, L. G. (1990b). A broadly based analysis of mathematical giftedness. *Intelligence, 14*, 327–355.
- Roff, M. (1941). A statistical study of the development of intelligence test performance. *Journal of Psychology, 11*, 371–386.
- Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology, 62*, 529–540.
- Tuddenham, R. D. (1948). Soldier intelligence in World Wars I and II. *American Psychologist, 3*, 54–56.

Chapter 9

*Theoretical Models for the Study of Intelligence**

Earl Hunt

The University of Washington

THE SUBJECT MATTER OF INTELLIGENCE

The term *intelligence* has been applied widely. Clearly it is something we want to have; who wouldn't want to be intelligent? But just what does "being intelligent" mean? Guilford (1967) identified over 100 separate human abilities that he regarded as part of intelligence, and he confined himself solely to the intellectual realm. Some authors want to expand the term to encompass artistic and even athletic ability (Gardner, 1983).

This will not do. Scientific analyses have to be limited to manageable sets of variables. If we make "intelligence" coterminous with "all the nice abilities a human might have" the topic will be so incoherent that it will never be understood. In this chapter I will try to establish a reasonably important, but not all-encompassing, sphere of human activity, and then try to define intelligence within that sphere.

*The preparation of this chapter has been partially supported by the Office of Naval Research, Contract N00014-86-000065 to the University of Washington; Earl Hunt principal investigator. The opinions expressed are solely those of the author.

Much of human activity is rational. Rational activity occurs when people decide to attain a goal and choose actions intended to reach it. Problem solving in algebra is an excellent example; people who solve algebraic problems do not write characters randomly, they plan their actions. They are fully aware of the fact that they are trying to solve an algebra problem, and, at some level, they are aware of how they are trying to solve it. To take a less academic example, when stockbrokers choose a stock or doctors recommend a treatment they are prepared to justify their recommendations.

Acts of rational thought may or may not be successful. A person who offers an invalid proof is thinking just as much as one who offers a valid one, and stockbrokers who recommend bad purchases are thinking as much as their more foresightful colleagues. They just are not thinking as well. When different individuals consistently display different levels of cognitive competence we say that they differ in *intelligence*.

This term has come to carry a good deal of excess emotional baggage, which is unfortunate. The term will be used here solely as a shorthand for “consistent individual differences in cognitive competence.” I will further limit the term by admitting that *consistent* can mean consistent over time, but not necessarily over domains. I, and many other people, can imagine stockbrokers who are poor mathematicians and vice versa.

As I shall use the term, intelligence refers to the quality of a person’s goal-directed problem solving. I will not address the issue of why, in some larger scheme, the person has chosen his or her particular goals. In addition, there will be no discussion of whether or not the person’s goals fit into some social structure where appropriate goals are chosen by another person. Although we might not admire the goals of, say, a bank robber, we could still admire a clever way of breaking into a vault.

These restrictions clearly rule out a lot of human mental behavior. There is no place for love, although there may be one for Machiavellian schemes for obtaining sex and wealth. More seriously, it is easy to create borderline cases. There probably was something to Freud’s claim that we are not always aware of the goals we are pursuing. It is certainly true that we are not always aware of the restrictions that we place on problem solving. It is also true that the success of intellectual efforts often depends upon others’ evaluations of the problem solver’s goal. Most professors at our great research universities make less money than the football coach. This does not necessarily mean that the coach is more intelligent than the professors.

It should be clear from these examples that I want to study an individual’s problem-solving ability, more or less in isolation from the social setting in which that ability is exercised. It can be argued that this is a mistake, because intellectual competence is always exercised within a particular social context. Several advocates of “situated cognition” have taken this approach. I disagree and believe that it is sensible to study the mental capacities of a single indi-

vidual, in isolation. In order to understand the person's total behavior, you may have to go beyond such studies, but the studies themselves are reasonable. Let me offer an analogy.

Suppose that instead of studying the behavior of humans we were studying the behavior of aircraft. We would soon find that the behavior of an aircraft depends on two different modular subsystems: the aircraft and the aviator. The total behavior of an aircraft in flight cannot be understood unless both are understood. Nevertheless, aeronautical engineering and human factors are separate disciplines. They can each be studied on their own, and the interface between them is also worth studying. Now let us look at this more formally.

A scientific analysis of any topic can only proceed if the topic approximates what mathematicians call a *closed system*. Formally, a closed system is a set of *observed variables* $V = \{v_i\}$ $i, = 1..K$ such that any one variable can be expressed as a function of the others, that is:

$$v_i = f(v_1...v_j...v_K) \tag{1}$$

A system is open to the extent that there are *unobserved* variables $X = \{x-w\}$ $w = 1...L$ not in V , such that some (possibly all) of the v_i 's are also functions of the $x-w$'s;

$$v_i = f(v_1...v_j...v_K, x_1...x_L). \tag{2}$$

Clearly if the unobserved variables exert a major influence on the observed ones we will not be able to obtain orderly data, and therefore any scientific endeavor is defeated to start with.

My contention is that the only way that we can define a manageable system for individual differences in cognitive competence is to restrict our attention to conscious, rational problem solving, being attempted by an individual. In fact, this area itself will have to be broken into subfields. If we try to develop a grand theory of social interactions and motivation at the same time, we will simply get nowhere because we are studying open systems, and science cannot do that.

The decision to restrict our attention to closed systems is not trivial. The decision is valid only if an individual's cognition can be treated as a module in a complex network of modules that includes other individuals, the physical environment, history, and so forth. In terms of the mathematical formalisms just introduced, it assumes that we can profitably study the relationships between the observed variables conditional upon certain (ranges of) values of the unobserved variables. The effects of changes in the unobserved X variables (e.g., changes in social reward systems) are treated as interface problems between the individual and the environment. As in the aircraft example, we do not want to confuse the study of the interface with study of the system.

THE CRITERIA FOR SCIENTIFIC THEORIES

Just what does it mean to say that a theory of intelligence should be a scientific theory? There has been rather a lot of debate about just what a scientific theory is. At the risk of sounding naive, I shall now attempt a small contribution to the topic.

Borrowing a term from modern cognitive psychology, a theory is a mental model that people use to deal with the phenomena at hand, here individual differences in mental competence. Mental models are handy because they let a thinker understand variations in the world by making manipulations in his or her internal representation of the world. This means that a theory is a tool, just as much as a hammer is a tool. In order to be of any use to the person holding the theory, the theory has to be comprehensible—so that it can be managed at all—and the theory has to produce information that the person can use to achieve his or her own goals.

These criteria are somewhat different from those conventionally taught in logic of science courses. They make theory personalistic; a theory that is comprehensible to one person may not be comprehensible to another, and a theory that is useful for one set of goals may not be useful for another. I think that theories have to be personalistic in this sense. More strongly, I believe that a certain amount of confusion has resulted because people have argued that theories of intelligence that are appropriate for some individuals and purposes are appropriate for all individuals and purposes.

Scientific theories must be objective. They must be publicly stated, in such a way that any appropriately trained individual can evaluate them as a model for some objectively defined behavior. Going back to the formalisms introduced a few paragraphs ago, a scientific theory is a closed set of *theoretical* variables, $Y = \{y_v\} \ v = 1 \dots M$, that are interrelated, that is:

$$y_v = g(y_1 \dots y_M) \quad (3)$$

In addition, in order to be objective a theory has to be accompanied by a clear-cut measurement procedure relating the y 's to the observable v 's. Thus for every v_i there must be a predicted value

$$v'_i = h(y_1 \dots y_M) \quad (4)$$

that it can be compared to. The closeness of this fit determines the accuracy of the theory.

Note that objectivity, as defined earlier, does not in any way conflict with my emphasis on personalistic theories. For me, for instance, modern theories of cosmology are not a satisfactory personalistic model of the world because I do not understand them. I understand the procedure required to comprehend them

(i.e., study advanced physics for several years) and I will take on faith the statement that the measures of cosmological theory can be defined objectively. I will also take on faith the assurance that cosmologists can, for instance, predict the bending of light around Jupiter more accurately than can be done by Newtonian optics. In my mundane world, I have no utility for this level of measurement; Newtonian optics will do just fine. Cosmology is objective but, for me, it is not personalistic.

Finally, if a theory is to have any content it must be *pragmatically true* rather than *necessarily true*. A pragmatically true statement is one that is capable of disproof by some conceivable observation, while a necessarily true theory is one that makes only tautologically true statements about observables.

This last point may be so obvious that it is hardly worth saying, but unfortunately the history of the study of intelligence contains more than a few examples of necessarily true theories. For instance, a theory of insight that states that insightful problem solvers retrieve the necessary background information to solve a problem is necessarily true and tells us nothing. References have been suppressed to protect the guilty.

TYPES OF THEORIES

The framework just described allows us to generate two different classes of scientific theory: *descriptive* and *causal* theories.

Descriptive theories are stated in terms of a relatively small set of relationships, called *laws*, that constrain the observables to certain sets of values. No causal mechanism is presented to explain why these laws hold, they just do.

The nature of gravity is explained this way. Students are told that gravity is a relationship between objects that depends on the masses of the objects involved and their distances from each other. No explanation is given of why gravitational forces exist, they just do.

Summarizing laws can be found in psychology. B. F. Skinner's formulation of behaviorism is the most obvious example. Another good example can be found right in the study of intelligence. Factor analytic theories are descriptive theories. To illustrate, consider Thurstone's (1938) famous principal factors model.¹ Thurstone asserted that data resulting from measurement of N individuals on K measures of mental competence could be summarized by stating positions for each of the N individuals on $M < K$ factors. He also provided a method for going from observable to factors and back again. The model is objective and, in a wide range of situations, comprehensive and useful. It is a collection of laws, rather than a collection of explanatory statements.

¹ I could have used any other factor analytic model in this example. Thurstone's was chosen solely because it seemed easiest to describe.

Causal theories go beyond descriptive theories by providing an explanation for observations in terms of some mechanism that forces a relationship to exist. Often causal relationships will be defined in terms of a more basic theory that is descriptive at its own level. For instance, ballistics treat gravity as a causal force determining the flight path of an object, although within the framework of Newtonian mechanics, gravitational relations are descriptive.

There are similar theories in psychology. Genetic models of intelligence imply certain statistical relationships between degree of family relationship and performance on cognitive tests. The direction of causation is clear. Genetic models themselves may be derived from causal theories in molecular biology, but at the Mendelian level, genetics is descriptive.

There is a third class of "intellectual objects" that are sometimes referred to as theories, but that I refer to as *world views*. A world view is a way of generating specific theoretical statements. Three examples from the study of intelligence should illustrate what I mean. Factor analytic theories of intelligence begin with the statement that an individual's performance on a test battery will be represented as a vector of real numbers. Specific theories are then generated by stating restrictions on the relationships between vectors (cf. the characterization of Thurstone's theory). Let me call this the *Euclidean world view*. It is a particular way of representing the data to be explained. It is not a theory itself, but it can be used as the starting point for generating theories.

A similar remark can be made about Sternberg's (1985, 1988) highly publicized Triarchial "Theory" (quotation marks very much mine). The triarchial theory defines three aspects of cognitive behavior: performing a cognitive act, monitoring one's actions (metacomponents) and acquiring new problem-solving skills. The triarchial theory fits into previous work (Sternberg, 1977) in which Sternberg has argued for a *componential world view* in which different types of cognitive actions are isolated and treated as separate topics of investigation, perhaps using different theoretical mechanisms. Here it is sufficient to note that this is a world view about how to represent cognitive behavior. It may or may not be useful. It is clearly different from the Euclidean view.

In the following section I shall offer what I will call a *computational world view*. To foreshadow, I shall argue that cognition is best thought of as a computation; that is the manipulation of a physical symbol system that is supposed to represent external reality. This idea is originally due to the pioneering work of Allen Newell and Herbert Simon (Newell & Simon, 1972). I then consider different aspects of a cognitive computational system, and argue that different theories of intelligence are needed to account for individual differences in each aspect of cognition.

World views are sometimes looked upon as almost religious belief systems. They cannot be verified or disproved, you either have to believe them or not. I am going to take a somewhat different approach. I believe that world views can be accepted or rejected, but not directly. Instead, a world view rises or falls

on the merits of the theories that it generates. Does a world view generate theories that meet the criteria enumerated for them? This is a matter of consensus, rather than something to be determined by a crucial test of a hypothesis.

This concludes my somewhat lengthy preamble. Next, let us turn to considering some specific aspects of the study of intelligence.

A THEORY OF COGNITION AND INTELLIGENCE

Scope of the Theory

This section will present a world view that attempts to unite theories of intelligence and cognition. In this world view intelligence is viewed as the product of cognition rather than the other way around. I object to a description of a person's cognition that says "X thinks that way because X is (or is not) intelligent." That is putting the horse before the cart. X is intelligent because of the way X thinks. Therefore a theory of intelligence should explain how variations in the parameters of cognition produce individual differences in mental competence.

My approach will emphasize causal relationships, explained in terms of psychological theory. Thus phrases like "genetic constitution causes intelligence" have no place in this approach (although they may well be true) because they are statements about how biological phenomena produce psychological phenomena, rather than statements about how one psychological phenomenon produces another.

Figure 1 shows the argument in diagrammatic form. Biological phenomena establish a basic cognitive capacity. This capacity is the topic of theories of cognition. Social experiences interact with the basic cognitive capacity to produce another.

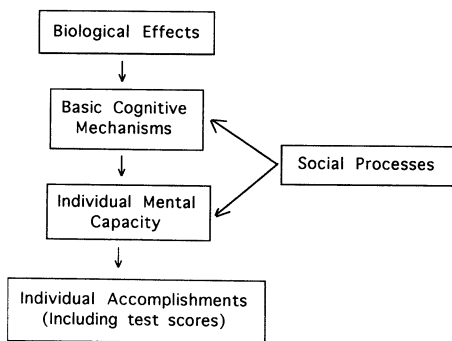


Figure 1. The relationship between biological and social influences, information processing, and demonstrated mental competence.

duce individual mental competence (i.e. intelligence). Mental competence, in turn, interacts with the social environment to produce indications of individual ability, including, but certainly not limited to, scores on intelligence tests.

A Model for Cognition

Since the 1950s, virtually all psychologists have accepted the proposition that thinking is a species of computation, and that the concepts applicable to computing systems are appropriate for the analysis of human thinking.² A particular model of computing, called the blackboard architecture model, has proven to be particularly useful (Hunt, in press; Hunt & Lansman, 1986; Klahr, Langley, & Neches, 1987). The basic idea behind the model can be grasped by a metaphor, which is portrayed in Figure 2. Imagine a stage—called working memory—in which all active thought takes place. This stage is divided into at least three smaller stages: auditory working memory, visual working memory, and an abstract memory for reasoning. These stages are where “conscious thought” takes place. In more computational terms, the temporary data structures required to solve the problem currently being worked on are stored in the various working memories.

The stages are viewed, if you will, by the contents of a very large long-term memory. The long-term memory contains rules for problem solving. These can be very simple, as in the rule “When driving and you see a red light and stop.” Note that this rule contains two parts, a pattern and an action to be taken when an example of a pattern is recognized in the appropriate part of long-term memory. Simple rules like this are called *productions*. There is an extensive literature on their use in psychological modeling, so there would be little sense in repeating the argument here.

Productions are themselves organized into larger units, called *schemata*. Just what a schema is has been the topic of a great deal of debate. For our purposes, a schema can be thought of as an organization of simple problem-solving steps (the productions) into an orchestrated but specialized problem-solving method. Here a metaphor to the use of a schema “outside the head” may be useful. The Internal Revenue Service’s famous IRS 1040 form is a schema. It tells you how to compute your income tax, in step-by-step fashion. Most of the steps are defined by the production rules for ordinary arithmetic.

The basic assumption behind this model of thinking (and that of many other people’s) is that the behaviors we call thinking depend very heavily on

²This basic proposition has been questioned by proponents of Connectionist theories of cognition, who wish to tie human thought directly to mathematical models of brain systems (Rumelhart, McClelland, & Hinton, 1986). Although this approach may be useful in the future, at present it has not been developed to the point at which it could reasonably be used to develop a theory of intelligence. See Hunt (in press) for a further discussion.

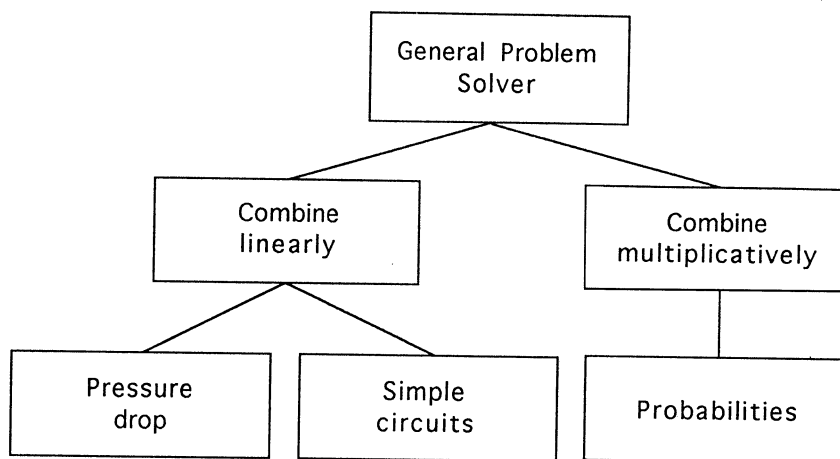


Figure 2. A hierarchial organization of problem-solving schema. This was suggested by Larkin's FERMI model, and subsequently amplified as a result of the works my colleagues and I have done on the psychology of physics instruction.

schematic reasoning. Such reasoning makes two qualitatively different demands upon the mind. The mind must have access to appropriate information, and the mind must have an adequate information-processing capacity, defined independently of the schemata themselves, to execute the computational steps the schemata demand.

To see this, let us go back to the IRS example. Form 1040 provides an economic way of calculating your income tax. If you do not believe this, imagine what would happen if you were provided with copies of the relevant tax law and nothing else! However, the form clearly assumes that you have certain information available: Form 1040 itself and a large array of bookkeeping documents. At a more basic level, it assumes that you know how to do arithmetic. In addition, the form implicitly assumes that you have certain information-processing capacities. It assumes you have access to a physical table or desktop sufficiently large to put the appropriate information down on it. It assumes that you have a device for writing down numbers, and that you have filed documentation for all your claims. Individual differences in these mechanistic processes of information handling, which are used for other schemata beside the tax forms, are an essential part of cognition.

The analogy is exact. We can think of a person's mental capacity as being determined by the schemata, the information available to fill in those schemata and by the person's capacities for manipulating information in the abstract.

To make things more confusing, “intelligence” (i.e., individual mental capacity) is not determined by any simple summation of schematic knowledge, information, and information handling. Some schemata can only be executed by people who have substantial information-processing capacities. Others determine the way in which different pieces of information are seen as relevant to each other. And finally, schematic reasoning can increase the efficiency of information processing capacity.

Hierarchies of Schematic Reasoning: An Example from Physics

There is a tendency to fall into the trap of saying that we “use schematic reasoning” and then let it go at that. This is not very informative. In the extreme, we have all acquired “automated” reasoning systems that are so overlearned that they act almost like conditioned reflexes. At the other extreme, we have developed schemata for reasoning that can solve a wide class of problems.

Stopping your car at a stop sign is a good example. This is something we do virtually automatically. We also clearly use schemata that are not automated, but certainly do guide our actions. To continue with the driving example, driving over familiar routes, becomes schematized, to the point that it is rather easy to take the normal turns, even when you intended to deviate from the route on that particular day. We also have generalized schemata for finding street addresses in new locations.

Our minds seem to be organized into a lattice of schemata of varying degrees of generality. To illustrate this in a somewhat more academic setting than driving, let us consider a computer program developed by Jill Larkin and her colleagues (Larkin, Reif, Carbonell, & Gugliotta, 1988) to simulate reasoning in elementary physics. The program, called FERMI for a variety of reasons, could solve problems in circuit electricity and hydrostatics. As will be seen, other areas of physics could be added rather easily.

FERMI contains a lattice of reasoning systems that operate at varying levels of generality. An example is shown in Figure 2, which illustrates a fraction of the FERMI scheme. At the lowest level, FERMI has schemata that are highly specific to the problem being worked on, such as a schema for solving circuits with single resistors, or pressure differential problems in which one object is over another in a uniform media. At a higher level, FERMI contains schemata for combining quantities linearly. Thus the pressure drop problem shown in Figure 3 can be solved by FERMI by combining its simple schema and its linear reasoning schema. Finally, at the highest level FERMI contains a general problem-solving schema, modeled on Newell and Simon’s (1972) GPS algorithm, that can be used to plan its attack on a variety of problems.

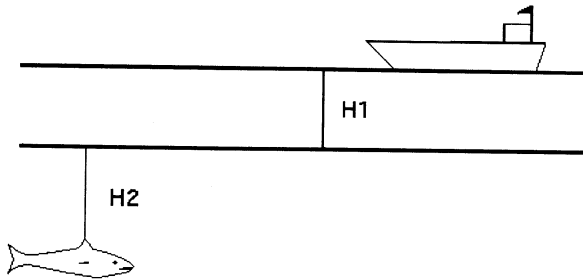


Figure 3. A pressure drop problem. What is the difference in pressure between the pressure on the ship and the pressure on the fish? The problem can be solved by applying hierarchical schematic reasoning.

The FERMI model could easily be extended to other domains. In fact, this is shown in Figure 2, which lists probability problems as a subset of the schema that use multiplicative combination rules. Probability theory was not part of, but clearly would fit into it.

Within its range of competencies FERMI was presented as a theory of perfect reasoning, rather than as a psychological model. To make theory a psychological model, and certainly in order to understand intelligence, we have to admit the fact that the schemata people have are not all as correct as the schemata nor are they that neatly related to each other. People may use schemata beyond their range of application, or they may fail to have correct schemata at all.

Indeed, there is considerable data from the study of physics indicating that much of the reasoning of elementary physics students is guided by what we could call "inconsistent" or "limited" reasoning. We see an example of this in the history of physics. Aristotle believed that objects had natural states, and that they tended to return to them. The natural state of a heavy object is to be at rest, and therefore heavy moving objects return to their resting position. Although Newton's laws of motion are valid in all situations, Aristotle's schema does provide a guide for predicting effects in a friction-filled world.

The idea that people often rely on schemata that have limited validity ties into a second important point about schematic reasoning and indeed, about reasoning within any blackboard architecture model. Localized schemata triggered by concrete events, are favored over general schemata. There is a good example in carpentry; the "3-4-5 rule." This rule states that if you want to construct a triangular joint the relation between the shortest side of the right angle to the other side and to the diagonal should be 3:4:5. This works for, say, building a table. It is a special case of the Pythagorean theorem; $C^2 = A^2 + B^2$. The

Pythagorean theorem is itself a special case of Euclid's laws of geometry. But a carpenter does not want to have to rederive the Pythagorean theorem every time a triangular joint is needed. Remembering the 3:4:5 relationship is much easier.

If you could predict the world that you were going to live in there would be no point in having generalized reasoning schemata. It would be much more efficient to have highly efficient, localized schemata that worked in specialized situations. Indeed, there is considerable evidence showing that this is what people do. The implication relationship of logic, for instance, seems to be understood by most people in terms of locally valid specializations rather than as a general schema (Cheng & Holyoak, 1985). Attempts to train logical reasoning are more effective if they build on the specialized schema, rather than if they attempt to teach logic in the abstract (Cheng, Holyoak, Nisbett, & Oliver, 1986).

To summarize, I propose that cognition is guided by a hierarchy of specialized and general reasoning schemata. Individual differences in cognition are a product of changes in the parameters of this hierarchical system. "Parameter" has to be understood as a very general term. It can refer either to the schemata that people possess, the factual knowledge that they use when they activate the schemata, or to the basic information-processing capacities that are required to activate the schemata. The link between cognition and education is provided by the schemata, while the link between cognition and biology is provided by the information-processing capacities. However, the linkage is not a matter of simply connecting systems together. Schematic problem solving is permitted by the information-processing system, and it determines how the information-processing system is used. Therefore it is not possible to say, in general, how well person *Y* will do on task *X*, given *Y*'s intelligence. We have to have a model of the cognitive demands of task *X* and a fairly sophisticated view of the mental capacity of person *Y*.

THE RELATION BETWEEN A SCHEMATIC REASONING MODEL AND THE CLASSIC FINDINGS ON INTELLIGENCE

There are certain "classic" findings that any theory of intelligence must deal with. In this section I will consider how these findings can be explained within the frame of a theory of schema-controlled cognition.

The Distinction Between Verbal Intelligence, Visual-Spatial Intelligence, and Reasoning

Theories of cognition and intelligence both differentiate between visual-spatial and verbal reasoning. Theories of intelligence do so on statistical grounds. The correlations between tests of verbal reasoning and visual-spatial reasoning are

substantially lower than the correlations between different tests of reasoning within each of these two domains. Cognitive psychology bases its distinction on rather more direct evidence. The two domains exhibit minimal interference. It is possible to perform verbal and visual-spatial tasks simultaneously, although it is quite difficult to perform two verbal or two visual-spatial tasks at the same time (Brooks, 1968; Wickens, 1980). There is also a great deal of neuropsychological evidence showing that verbal and visual-spatial reasoning can be selectively damaged by brain injuries (Shallice, 1988).

Blackboard architecture theories (of which schematic reasoning theories are one variety) handle this by assuming that working memory is divided into three separate parts: an auditory store that quite literally holds an "echo" of words that are either heard or generated internally, a more generalized working memory for reasoning, without a strong auditory component, and a specialized visual-spatial "scratch pad" storage area (Baddeley, 1986; Hunt, in press). Within each of these areas different information-processing mechanisms are required and, not surprisingly, individuals who are skilled in one type of processing may not be skilled in another.

This observation is fortified by a very large body of evidence showing that there are separate but correlated "abilities" (i.e., dimensions of individual differences) within each domain. Studies of spatial-visual reasoning show that at least four different abilities exist: the ability to keep a number of figures in memory at the same time, the ability to rotate objects "in the mind's eye," an ability to make judgments of motion, and an ability keep track of one's position in very large space (Eliot, 1987; Hunt, Pellegrino, Frick, Farr, & Alderton, 1988) These abilities are distinct, but related (i.e., the statistical relationship is neither zero nor one).

Much the same can be said about verbal reasoning. It appears to be broken down into two correlated components: roughly, the ability to retrieve relevant information and the ability to manipulate verbal expressions in working memory (Hunt, 1987; Palmer, MacLeod, Hunt, & Davidson, 1985). As in the case of the various spatial abilities, the verbal abilities are distinct, but related.

These results are well known. They also fit in well with the cognitive theory espoused in the previous sections. The various tasks used to measure different aspects of spatial-visual and verbal ability all call on several information-processing capacities (e.g., the ability to hold an image in working memory), but the mix of capacities required varies from task to task. The situation is analogous to a carpenter who uses different tools to build different things. In the mind, our tools appear to fall into two broad categories: tools for manipulating verbal or visual-spatial working memory. There seem to be individual differences in our facility for dealing with each tool.

Statistically, deductive reasoning ability stands somewhat apart from both verbal and spatial-visual reasoning. As is well known, abstract deductive reasoning is extremely difficult. Even very bright university students who are not

mathematicians have trouble with conditional reasoning in the abstract (Wason & Johnson-Laird, 1972). Similar difficulties are found with arithmetic reasoning. Paradoxically, though, anthropologists have observed that people who are not at all skilled at formal mathematics do reasonably well when mathematical problems are presented to them in a familiar, everyday context (Lave, 1988). Grocery shoppers, for instance, perform quite well in the context of shopping, but may do poorly on an abstract test of arithmetical reasoning.

This finding fits well with the schema-theoretic approach. Recall my earlier point that if you can anticipate all the problems you are going to face, special-purpose schemata are actually more useful than (computationally intense) general-purpose schemata. Yet when academics evaluate reasoning ability, they are interested in general reasoning. This does not mean that the academics are wrong; general reasoning is a very good thing in the appropriate context. The demand for general reasoning ability may not be as widespread as a university general studies curriculum would lead you to believe.

The Gc-Gf Distinction

There is a great deal of research supporting Cattell (1971) and Horn's (1985) distinction between fluid and crystallized intelligence—Gf and Gc. Roughly speaking, fluid intelligence is the ability to reason about unfamiliar problems, while crystallized intelligence is the ability to apply, in Cattell's terms, culturally approved problem-solving methods to previously encountered problems. This distinction clearly maps onto the distinction between specialized and general problem-solving schemata.

Horn and Cattell's argument for the Gf–Gc distinction was based on statistical evidence. A schematic theory of intelligence complements their observations by describing why the Gf–Gc distinction occurs. It also carries a further implication: that Gc will be relatively impervious to minor variations in information-processing capacity, whereas Gf will not. The reason is that Gc is a statistical manifestation of the exercise of highly efficient concrete schema, while Gf is a manifestation of the use of general but computationally intense problem-solving schema.

One of the classic findings in gerontological research is that Gf declines in the later adult years, while Gc does not. According to schematic reasoning theory this decline can come about for two reasons. One is simply the dominance of concrete over general schemata. As a person comes to acquire more and more specific schemata, there is less need to use general schemata. Horn has made the same point in several of his publications. He points out that as peo-

ple acquire Gc, they have less need for Gf. This is particularly true if one lives in a fairly stable cognitive environment.

This raises an interesting question. Why should Gf skills (generalized schema) atrophy as Gc abilities become paramount? Of course, one possibility is that they are forgotten. But what does it mean to forget? The simplest answer is that the skills are simply lost. Another possibility is that the skills are not lost, but the cues for using them are. For instance, the schemata required for algebra, a very general problem-solving method, seem to be retained for most of the adult lifespan (Baird & Hall, 1991). It would be interesting to know whether there is a decrease in the use of general problem-solving methods over the adult lifespan, conditional upon a demonstration that the abilities are present to be used, if the situation elicits them. At present such data are lacking.

There is another, somewhat less benign, reason for the decline of general reasoning skills. Cognitive gerontologists have documented a generalized slowing in information-processing capacity with increasing age. This is generally considered to indicate reduced efficiency in information processing (Cerella, 1985). Reduced information processing would impact most severely on the computation-heavy general reasoning processes than it would on the more efficient specialized schemata.

While the Gc-Gf distinction is usually made to illustrate the effects of aging, it is worth considering what it means for society. Not all societies value abstract thought. Concrete specialized reasoning is more efficient in situations where the same problems occur over and over again, and in situations where problems have to be solved quickly. Two quite different studies provide illustrations.

In the 1920s the great Soviet neuropsychologist Luria conducted some extensive ethnographic studies in Soviet Central Asia (Luria, 1976). His work in many ways foreshadowed more recent studies by American cognitive anthropologists. Luria found that in the more traditional societies people were decidedly deficient in the use of what I have called general problem-solving schemata. Indeed, there was antipathy and sometimes downright hostility to the idea that people should rely on reasoning rather than experience.

We do not have to go to exotic, non-Western lands to see this sort of reasoning. It occurs in our own society. Scribner (1984) reports an interesting and instructive example. She observed the reasoning processes used by diary workers when they loaded supplies for retail orders onto trucks. In theory, this is a problem in arithmetic. In practice, the more efficient workers visualized the movement of containers from opened cases onto the truck beds. The visualization strategy worked more quickly than the arithmetical one. Of course, the

visualization strategy would have to be relearned if the containers changed in size, but so long as this did not happen visualization was more rapid and as accurate as the general arithmetic strategy.

My point is that general reasoning skills are devices that let a culture go about its business. There is no difference, in principle, between an intellectual tool and a physical tool. In many settings general reasoning is a very useful tool, but it is not an essential one. People who are raised in a society that does not value generalized reasoning probably will not develop the skill. This does not mean that they are mentally inferior, in a biological sense. It does mean that they lack certain mental competencies.

THE STATISTICAL DISTRIBUTION OF INTELLIGENCE

This section deals with three findings from statistical studies of intelligence. Two of the findings are somewhat paradoxical. The third, although long suspected in some quarters, has only recently been well documented. I will argue that they can be tied together by schematic theories of intelligence.

The statistical existence of general intelligence (g) is well known. Virtually all tests of cognitive competence bear some relation to each other. Since there is clearly a common statistical component to tests with widely different content, there is a temptation to conclude that variations in general intelligence are due to variations in a biological component of intelligence (Eysenck, 1986).

A biological cause for intelligence would have to act on information-processing capacity rather than knowledge. We do not inherit knowledge about arithmetic, but we may inherit the information-processing capacity required to learn arithmetic. Paradoxically, though, the best statistical markers for g are not tests of information-processing capacity. General intelligence is best evaluated by complex reasoning tests, such as the Raven Matrix test (Snow, Kyllonen, & Marshalek, 1984).

The mathematical model behind psychometric theories of intelligence generally assumes that test scores are multivariate and normally distributed. This is probably not the case. Detterman and Daniel (1989), in a very carefully done study, showed that the correlations between conventional tests of intelligence are higher at the lower end of the general competence scale.

A good way to visualize these results is to consider what they mean for a graphic illustration of "the distribution of intelligence." Let us assume that intelligence can be measured by two separate abilities (e.g., verbal and visual-spatial ability). Figure 4 plots the distribution of measures on two such variables, under two different assumptions. The first, in panel (a), is that the two measures are distributed normally, with some unknown but positive correlation between them. The key point is that the correlation is assumed to be the same at all levels of each variable. A less statistical way of saying this is that low

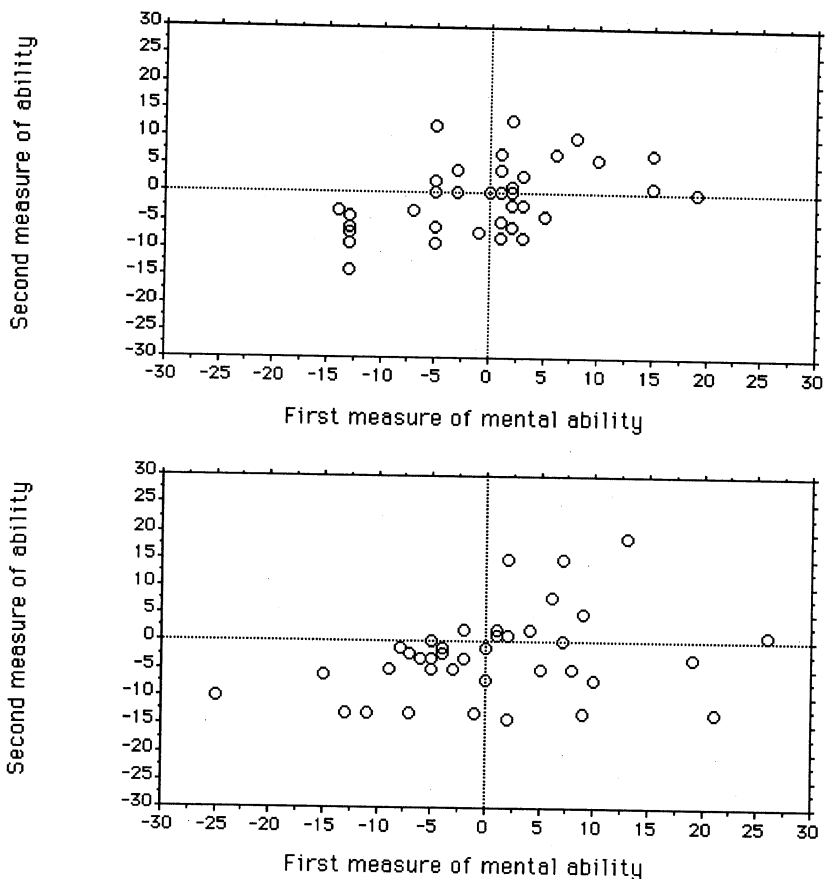


Figure 4. Distributions of scores on two different measures of mental capacity that would be obtained if different models of intelligence were true. The upper panel shows the scattergram if the standard multivariate model were true. Note that the points fall in a roughly elliptical pattern. The lower panel shows the scattergram that would result if poor performance on the first test accurately predicts poor performance on the second, but good performance is not predictive of good performance.

scores on one variable predict low scores on the other, and that high scores on one variable predict high scores on the other. As the figure shows, the data points lie in a roughly elliptical pattern, with very few points representing (high-low) or (low-high) pairs.

To relate to theories of intelligence, this is the pattern that would be expected if all test scores were determined by something like Spearman's (1927)

classic model of general and specific intelligences and there was a nonnegligible general intelligence factor.

Panel (b) shows the form of the distributions suggested by Detterman and Daniel's (1989) results. This data set was generated by assuming that low scores on measure 1 were fairly accurate predictors of low scores on measure 2, and that predictability decreased progressively as the level of measure 1 scores increased. Another way of saying this is that low scores on measure 1 go with low scores on measure 2, but high scores might go with anything.

Instead of being elliptical, the distribution takes on a shape something like an ice cream cone. What might produce such distribution?

Recall that schematic reasoning theory postulates two causes for Gf and Gc: individual differences in the possession of generalized reasoning schemata and individual differences in the possession of the information-processing capacities required to use them. At moderate to high levels of information-processing capacity (i.e., where most of us are); the possession of schemata determines reasoning ability. Since schemata possession is determined by social experiences, we would anticipate people displaying reasoning abilities in some fields and not in others.

At the low end of information-processing capacity, things change. Consider an extreme. Suppose that someone had no ability to hold temporary information in working memory. The person would be like my hypothetical taxpayer but without a desk. The taxpayer could not complete IRS 1040 or any other form. In the more psychological case, absence of any working memory capacity would make it difficult to use any sort of schematic reasoning outside of highly stereotyped, automatic behavior.

Biological Influences on Thought

One of the oldest arguments in psychology is the debate over the role of biological differences as determinants of individual differences in cognition. There are several subsidiary questions: How much of individual differences in psychology are due to biology? What biological variables are important? How do they exert their effects? The first question is one that is relevant to populations

³An example of this sort of question is "How much of the variance in IQ scores is associated with biological measures?" The answer depends upon the population. For instance, Herrnstein (1977) pointed out that if all children were raised in identical environments the heritability coefficients for that population would increase. This statistical fact, for some reason, provoked extreme anguish and even charges of racism. A "how much" question could be posed for an individual. For instance, during most of his early adulthood, King Henry VIII of England was an adept politician and, contrary to popular belief, apparently had a quite stable marriage to Catherine of Aragon. His political judgment became impulsive, and his wives became legend, after he received a severe blow to the head while jousting. It is an interesting question how much of Henry's impulsive behavior, including all the wives, was due to forebrain injury. This is a "how much" question, but it is not a question about variances across individuals.

rather than individuals.³ Therefore, it will not be discussed here. The other questions are questions that can be addressed within the context of a schematic theory of intelligence.

In discussing biological effects, one overarching principle has to be kept in mind. Biological variables influence information processing. They do not directly influence the content of the schemata that a person acquires. Nevertheless, the presence of biological effects may be best revealed by looking at complex rather than simple tests of elementary information.

The argument has been put well in a theoretical paper by Meyerson, Hale, Wagstaff, Smith, and Poon (1990). They began with the previously observed fact that there is a linear relationship between the time that it takes old and young individuals to do a task. That is

$$RT_{\text{old}} = A + B(RT_{\text{young}}) \tag{5}$$

where RT stands for reaction times of old or young people on some task, and A and B are positive real values. Now suppose that a task can be characterized by k stages, each taking the same approximate amount of time. Then for any group, x , that takes time T_x to complete a stage, a task will be completed in time

$$RT_x(k \text{ stage task}) = A + kT_x \tag{6}$$

Meyerson et al. (1990) argue that as people age, their information processing within each stage becomes unreliable. Therefore the time required to pass successfully through a stage (i.e., with reliable results) increases with age, because some form of recomputation is required to extract signals from noise. Now consider the problem of “achieving reliable results” in discriminating between data sets obtained from young and old subjects. The expected difference for a task with k stages is

$$E(RT_{\text{old}} - RT_{\text{young}}) = k(T_{\text{old}} - T_{\text{young}}), \tag{7}$$

which clearly increases with k . In more general terms, the more complex the cognitive task, the more likely the task is to show age differences, even though the “real” difference between old and young is in a simple act of information processing.

Very much the same argument can be applied to studies of the biological heritability of intelligence. Massive evidence has shown that intelligence test scores behave as if they were measures of an inheritable trait, with heritability factors on the order of .5 or greater (Bouchard, Lykken, McGue, Segal, &

Tellegen 1990; Plomin, 1988). This is somewhat puzzling on logical grounds, since clearly the behavior exhibited on intelligence tests cannot be inherited in the sense that eye color is inherited. No one acquires vocabulary knowledge as part of his or her genetic inheritance. A similar argument can be made for the sorts of complex visual-spatial and logical reasoning required to solve “culture-free” intelligence tests, which have sizable heritability coefficients.

The argument just given for age can be applied to the heritability data. The more complex a problem-solving procedure is—and the less problem solving has become schematized—the more sensitive behavior is to variations in the underlying information processing actions. Thus the best measurements of genetic variation may involve complex tests, even though, on logical grounds alone, we know that the behaviors and knowledge needed to solve the test problems could not possibly be genetically acquired. The reason is that studies of the genetic basis of intelligence have been based on descriptive rather than causal theories.

CONCLUDING REMARKS

One of the great mistakes of educators in psychology was to split the study of intelligence—individual differences in cognition—apart from the study of cognition itself. It is easy to understand why the split occurred during the heyday of behaviorism; it is less clear why it has been maintained during the era of cognitive psychology.

In this chapter, I have argued that the theoretical structure of cognitive psychology provides a world view that can be used to structure findings about individual differences. A world view is not itself a model, but it can be used to generate testable models. As such, it is very, very useful.

I have said very little about testing (i.e., about the measurement of intelligence). The omission was intentional. According to the world view espoused here, testing is an engineering operation. It does not make sense to measure cognitive processes until you have some idea of which processes you wish to measure and why.

Testing is done for one of two purposes: either to determine what effect some independent variable has had upon mental competence, or in order to predict how an individual will respond to the cognitive demands of a specific environment. According to the cognitive viewpoint, one cannot do this by “testing intelligence.” Indeed, from the cognitive viewpoint, a study of, say, the effect of improved nutrition upon intelligence test scores is close to pointless. But this does not mean that the study is silly in the absolute. There are many contexts in which establishing the relationship would be an appropriate thing to

do. Measurement is an engineering operation. A measurement is good or bad depending upon whether or not it fulfills its purpose.

One of these purposes is to understand how some independent variable works. Since we seldom, if ever, choose our independent variables at random, it makes sense to consider how the independent variable in question is supposed to work, and to choose one's dependent variables accordingly.

A study on the psychopharmacology of cognition illustrates this point. MacLeod, Dekaban, and Hunt (1978) were interested in the relationship between phenobarbital and cognition. Their interest stemmed from some clinical observations. Epileptic patients who received phenobarbital to control seizures had reported that they felt "befuddled." Further inquiry suggested that they had trouble remembering things, especially in situations where events were moving rather quickly. Accordingly, MacLeod et al. (1978) studied the effect of the drug on the ability to retrieve information from short-term or long-term memory. They found that only short-term memory retrieval was affected.

The same sort of observation can be made about educational interventions. Minstrell and I have been studying the effects of a "cognitive science oriented" approach to teaching the concepts of physics areas (Hunt & Minstrell, in press). We are now trying to determine the extent to which instruction in physics generalizes to reasoning about topics in physics that were not part of the instruction, and even; further to analogous situations outside of physics itself, other investigation are involved in similar projects. For instance, Nisbett and his colleagues (Nisbett, 1992) have reported similar generalizations in the use of statistical reasoning beyond the area of original instruction.

I am afraid that too many psychologists would say that neither of these studies were studies of "intelligence" because they did not involve an intelligence test. I believe that the objection uses a particular measurement device to reify a concept. The studies cited were certainly studies of causes of individual difference in mental competence. If that is not intelligence, what is? Furthermore, the measurements used were chosen because they fit into a theory of how cognition is influenced by drugs, on the one hand, or the possession of schemata, on the other.

A second purpose for measuring intelligence is to predict how well a person is likely to do in some important setting (e.g. in school or in the workplace). Related to this purpose, there are some situations in which the intelligence measure may serve as a surrogate for real-world cognitive variables that are difficult or impossible to predict. Return to the hypothetical study of the relation between intelligence and nutrition. Such a study would be quite reasonable if the purpose was to assess the costs and benefits of, say, an enhanced nutritional program for young children. However, the purpose of the study would

be to let us make a more educated guess about how the children might do in the real world, not to learn about the mechanisms of individual differences in cognition.⁴

However, we should not make the mistake of thinking that theories of intelligence are solely of interest to academicians, while tests are the province of hard-headed, practical clinicians, educators, and industrial-organizational psychologists. Whenever possible, theory ought to be applied to situations in which tests are used as predictive devices. Now too much of present practice involves taking an omnibus measure of "intelligence" and correlating it with some equally omnibus (and often highly suspect) measures of criterion performance. From the viewpoint of cognitive psychology, it would make more sense to analyze the criterion situation and determine the sorts of cognitive demands it places on the individual. The next step is to determine how to evaluate a person's ability to respond to these demands, rather than to determine how they "think in general." This is precisely what Binet did for the French school system, and his test was a good one in that setting. His example should be followed more widely.

Of course, this means a good deal of work. In order to produce predictive tests, psychologists will have to learn as much about behavior in the criterion situation as they do about behavior in the test situation. But what is wrong doing a bit of human engineering? Indeed, it is rather arrogant of psychologists to claim that they can predict success in jobs X, Y, and Z unless they know rather well what jobs X, Y, and Z are. This is more than a call for obtaining validity coefficients. Validating a model of criterion performance ought to be a prerequisite for any test development project.

REFERENCE

- Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Oxford University Press.
- Bahrlick, H. P., & Hall, L. K. (1991). Lifetime maintenance of high school mathematics content. *Journal of Experimental Psychology: General*, *120*, 20–33.
- 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*, 229–236.
- Brooks, L. R. (1968). Spatial and verbal components in the act of recall. *Canadian Journal of Psychology*, *22*, 349–368.

⁴Since passions run high about such things, I hasten to add that in some situations there are excellent reasons for advocating nutritional programs, even if they have no effect whatsoever on children's cognition. This is certainly true if the children involved are malnourished. Making unhealthy children healthy is a good thing to do, and needs no further justification. A proposal to add vitamin supplements to the diets of normal, healthy children might well be scrutinized more closely.

- Carroll, J. B. (1982). The measurement of intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 29–122). Cambridge, UK: Cambridge University Press.
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston, MA: Houghton-Mifflin.
- Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin*, *98*, 67–98.
- Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*(4), 391–416.
- Cheng, P., Holyoak, K. J., Nisbett, R. E., & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology*, *18*, 293–328.
- Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest in low IQ groups. *Intelligence*, *13*, 349–360.
- Eliot, J. (1987). *Models of psychological space: Psychometric, developmental, and experimental approaches*. New York: Springer-Verlag.
- Eysenck, H. (1986). The theory of intelligence and the psychophysiology of cognition. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 3, pp. 1–34). Hillsdale, NJ: Erlbaum.
- Gardner, H. (1983). *Frames of Mind*. New York: Basic Books.
- Guilford, J. P. (1987). *The nature of human intelligence*. New York: McGraw-Hill.
- Herrnstein, R. J. (1977). *I. Q. in the meritocracy*. Boston, MA: Atlantic Monthly Press.
- Horn, J. L. (1985). Remodeling old models of intelligence. In B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurements, and applications* (pp. 267–300). New York: Wiley.
- Hunt, E. (1987). The next word on verbal ability. In P. A. Vernon (Ed.), *Speed of information processing and intelligence*. Norwood, NJ: Ablex.
- Hunt, E. (in press). *Thoughts on thought: A discussion of formal models of cognition*. Hillsdale, NJ: Erlbaum.
- Hunt, E., & Lansman, M. (1986). A unified model of attention and problem solving. *Psychological Review*, *93*, 446–461.
- Hunt, E., & Minstrell, J. (in press). A cognitive approach to the teaching of physics. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice*. Cambridge, MA: MIT Press/Bradford Books.
- Hunt, E., Pellegrino, J. W., Frick, R. W., Farr, S., & Alderton, D. (1988). The ability to reason about movement in the visual field. *Intelligence*, *12*, 77–100.
- Klahr, D., Langley, P., & Neches, R. (1987). *Production system models of learning and development*. Cambridge, MA: MIT Press.
- Larkin, J. H., Reif, F., Carbonell, J., & Gugliotta, A. (1988). A flexible expert reasoner with multi-domain inferencing. *Cognitive Science*, *12*(1), 101–138.
- Lave, J. (1988). *Cognition in practice*. Cambridge, UK: Cambridge University Press.
- Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations* (Lopez-Morillas & Salotaroff, Trans.). Cambridge, MA: Harvard University Press.
- MacLeod, C. M., Dekaban, K. A., & Hunt, E. (1978). Memory impairment in epileptic patients: Selective effects of phenobarbital concentration. *Science*, *202*, 1102–1104.

- Meyerson, J., Hale, S., Wagstaff, D., Smith, G., & Poon, L. (1990). The information loss model: A mathematical model of age-related cognitive slowing. *Psychological Review*, 97, 475–487.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Nisbett, R. E. (1992). *Rules for reasoning*. Hillsdale, NJ: Erlbaum.
- Palmer, J., MacLeod, C. M., Hunt, E., & Davidson, J. (1985). Information processing correlates of reading. *Journal of Memory and Language*, 24, 59–88.
- 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. 1–71) Hillsdale, NJ: Erlbaum.
- Rumelhart, D. E., McClelland, J. L., & Hinton, G. E. (1986). The appeal of parallel distributed processing. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing* (Vol. 1: Foundations, pp. 3–44). Cambridge MA: MIT Press.
- Scribner, S. (1984). Studying working intelligence. In B. Rogoff & J. Lave (Eds.), *Everyday cognition: Its development in social context*. Cambridge, MA: Harvard University Press
- Shallice, T. (1988). *Cognitive neuropsychology*. Cambridge, UK: Cambridge University Press.
- Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 47–104) Hillsdale, NJ: Erlbaum.
- Spearman, C. (1927). *The abilities of man*. London: Macmillan.
- Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Erlbaum.
- Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge, UK: Cambridge University Press.
- Sternberg, R. J. (1988). *The triarchic mind: A new theory of human intelligence*. New York: Penguin.
- Thurstone, L. L. (1939). *Primary mental abilities*. Chicago, IL: University of Chicago Press.
- Wason, P. C., & Johnson-Laird, P. N. (1972). *The psychology of reasoning: Structure and content*. Cambridge, MA: Harvard University Press.
- Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson & R. Pew (Eds.), *Attention and performance* (Vol. VIII). Hillsdale, NJ: Erlbaum.

Chapter 10

Phlogiston, Animal Magnetism, and Intelligence

Arthur R. Jensen

University of California, Berkeley

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

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

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

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

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

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

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

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

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

THE EMPIRICAL PHENOMENON OF VARIANCE IN MENTAL ABILITIES

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

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

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

FROM ABILITIES TO TESTS AND FACTORS

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

THE RELATIVE IMPORTANCE OF THE *G* FACTOR

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

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

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

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

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

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

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

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

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

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

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

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

EMPIRICAL FINDINGS GERMANE TO THE EXPLANATION OF *G*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

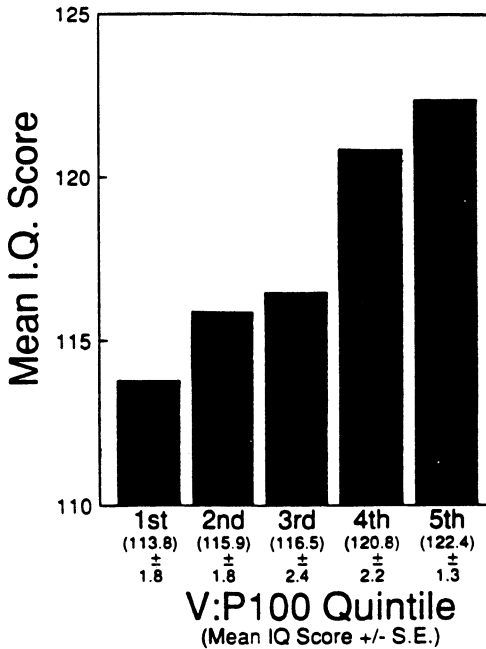


Figure 1. Distribution of mean IQ scores in V:P100 quintiles. The distribution of V:P100 values (i.e., the NCV based on the P100 latency) of the 147 students, from the lowest NCV (1.75 m/sec.) to the highest (2.22 m/sec.) was divided into quintiles. Quintile 1 contains the 20% of students with the values,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

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

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

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

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

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

Why the Apparent Ceiling on the RT- g Correlation?

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

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

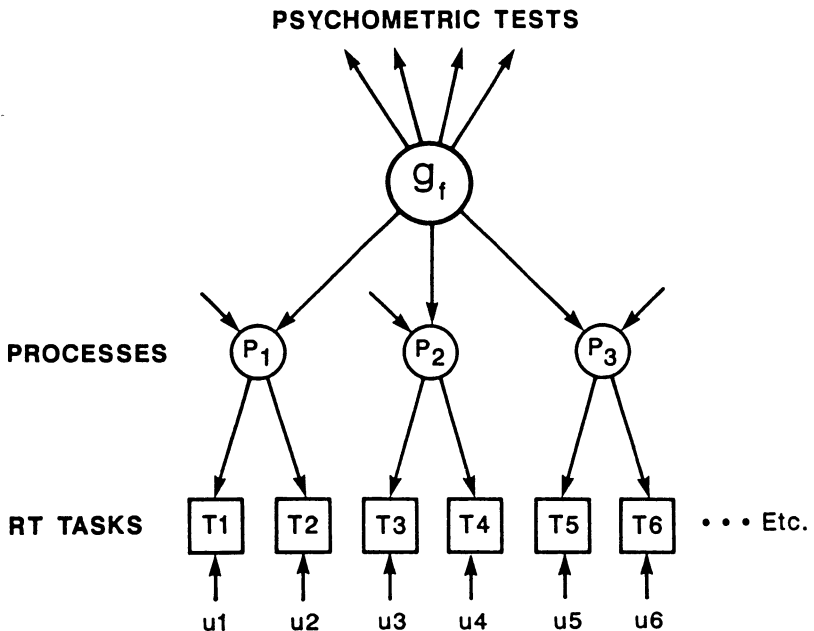


Figure 2. Representation of the factor structure of RT measures on a battery of diverse ECTs, in which some unspecified number information processes (P) tapped by the ECTs accounts for all of the variance in psychometric g .

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

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

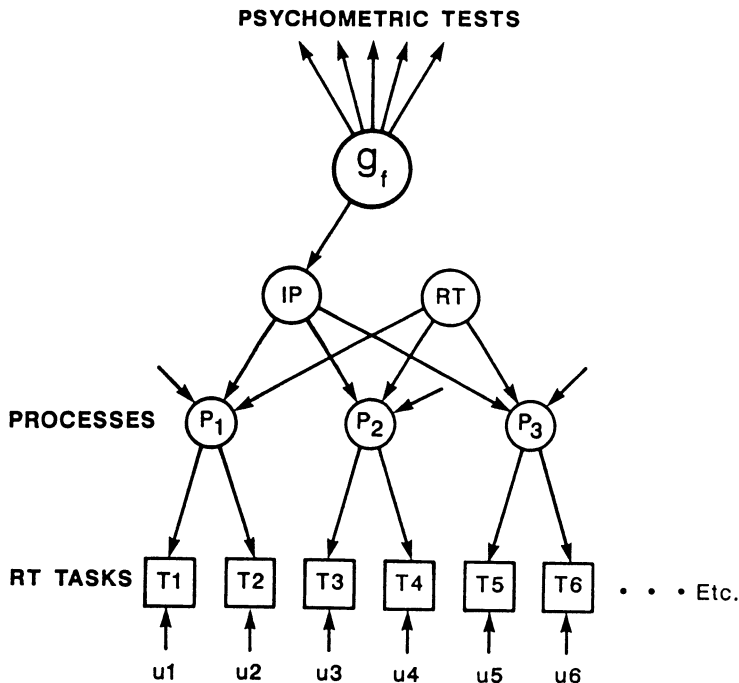


Figure 3. Representation of the factor structure of RT measures on various ECTs in which only part of the RT variance is associated with information processes (IP) and part of it is due to noncognitive factors. Only the cognitive or information processing (IP) part of the total RT variance is related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

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

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

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

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

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

A THEORETICAL FORMULATION OF *G*

Why Is Speed of Information Processing so Important?

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

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

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

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

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

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

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

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

DISCLAIMER

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

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

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

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

REFERENCES

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

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

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

Chapter 11

***Multiple Intelligences in
Multiple Contexts****

Mara Krechevsky

*Harvard Project Zero
Harvard Graduate School of Education*

Howard Gardner

*Harvard Project Zero
Harvard Graduate School of Education
and
Boston Veterans Administration Medical Center*

INTRODUCTION

Consider the following scenario: At a public elementary school in downtown Indianapolis, children take classes not just in reading, arithmetic, and social studies, but in music, dance, visual arts, computing, and Spanish. These latter

*The research reported in this article has been generously supported by grants from the William T. Grant Foundation, the James S. McDonnell Foundation, the Rockefeller Brothers Fund, the Spencer Foundation, and the Bernard Van Leer Foundation.

subjects are not considered “extra” or frills, but are given the same amount of support and time as the more standard curriculum. On a given day, a visitor to the school is as likely to find children engaged in learning a sequence of dance steps or composing a simple tune as puzzling out a word problem or answering questions on a social studies chapter. The children also work on projects related to schoolwide themes, which are subsequently videotaped and become part of a “video portfolio.” The portfolio documents a child’s development throughout her tenure at the school.

Next consider an elementary school in Gloucester, MA, where children in kindergarten engage in an unusually wide array of activities. In one classroom, children participate in activities related to the music and art areas; in another, social and math activities are provided; in the third, mechanical and movement activities are primary. Language development is encouraged throughout the curriculum. The areas for each class are chosen based on teacher interest and confidence in working in the domain. Many of the activities relate to a common theme, such as “community” or “time.” Parents, businesses, and local museums also serve as resources for the classroom. For example, as part of a community unit, children visit the local fisheries, and parents working in the fishing industry visit the classroom to talk about their experience.

Finally, consider a program for gifted children who are also disadvantaged. Rather than relying on traditional measures of intelligence which focus on language and logic to determine entry into the program, admission is based on performance on a wide range of measures. For instance, children are asked to tell stories, using a small storyboard; they take apart and put together mechanical objects; and they sing a song and match pitches played on a xylophone. Once children are identified as exhibiting strengths in one or more of these areas, the program provides experiences and activities which support the development of the different competencies that have been identified.

What do these three scenarios have in common? First, they each describe a program which both exposes children to and encourages development in a variety of areas. Second, in each scenario, children are working on projects and activities which are easily translated into products and experiences of real-world significance. Third, and most important for our present purposes, each one of the above scenarios is motivated in part by the theory of multiple intelligences put forth by Howard Gardner in 1983.

In this chapter, we describe a pluralistic theory of intelligence which suggests that our view of competence should be expanded to include domains beyond those typically examined in most intelligence tests. Although multiple intelligences theory as originally set forth located intelligence inside the heads of individuals, recently Gardner, among others, has begun to consider the intelligences and their combinations across a range of contexts. After summarizing the main components of the theory of multiple intelligences, we present our

revised theory of contextualized intelligence. Finally, we consider the implications of our theory for assessment and education.

In most Western cultures, intelligence has traditionally been considered a biological or psychological potential, situated inside the individual's head, which can be measured through standardized tests. While this line of thought can be traced as far back as Plato, a significant step in the measurement of intelligence was taken at the beginning of the century by the French psychologist Binet. Binet was given the task of identifying different degrees of mental retardation in French school children. He devised a set of test items, ranging from sensory discrimination to sentence completion, which became the prototype for subsequent scales for measuring mental ability (Sattler, 1988).

Although Binet believed that intelligence was too complex to capture with a single number, and did not think his scale should be used to rank students on a continuum of mental worth, the notion of intelligence as an entity with an independent existence readily gained prominence in this country (Gould, 1981). Terman at Stanford University adapted and normed the Binet-Simon scales for widespread administration. The revised Stanford-Binet Intelligence Scale, along with other IQ measures, is currently used in a variety of settings ranging from schools to the military. Although such tests have been shown to be reasonably reliable predictors of scholastic success, they do not yield impressive correlations with performances in the work world (Ceci, 1990; Jencks, 1972; Okagaki & Sternberg, 1990; Sternberg & Wagner, 1986). Indeed, most IQ items can only with difficulty be related to desirable real-world consequences, and the results of these tests have had little implication for designing and structuring educational curriculum.

THE THEORY OF MULTIPLE INTELLIGENCES

The theory of multiple intelligences (hereafter referred to as MI theory) challenges the notion that intelligence can be considered a general ability, which cuts across all domains of competence. Rather, Gardner argues for the existence of seven relatively autonomous intellectual competences, which inhere in varying degrees in all individuals: linguistic, musical, logical-mathematical, spatial, bodily-kinesthetic, interpersonal, and intrapersonal. Although there have long been competing views on whether or not intelligence could be considered one entity (Eysenck, 1981; Jensen, 1969; 1980; Spearman, 1923) or many (Guilford, 1967; Thorndike, Bregman, Cobb, & Woodyard, 1926; Thurstone, 1938), what is unique about Gardner's theory are the types of evidence he examines to derive his candidate intelligences.

According to Gardner, an intelligence can be defined as the capacity to solve problems or fashion products which are valued in one or more cultural settings. This definition reflects the idea that intelligence cannot be considered

apart from the *uses* to which it is put or the *values* of the surrounding community. Intellect cannot be conceptualized or assessed in the abstract. All intellectual activities take place in a particular domain or discipline. Indeed, as we shall maintain, a theory of context is needed to complement a theory of individual competence in order to describe fully and accurately the range of human intelligences.

To determine whether or not a particular capacity qualifies as a "human intelligence," Gardner took into account eight different criteria. The first source of evidence he considered was the potential isolation of an intellectual competence by brain damage. To the extent that an ability could be destroyed or spared by conditions of brain damage, in isolation from other abilities, its relative independence as an "intelligence" seemed likely. The specificity of certain kinds of cognitive functioning can be tied quite precisely to particular regions of the brain. For example, damage to the left frontal lobe produces difficulty in grammatical speech, while comprehension is relatively preserved. On the other hand, damage to the left temporal lobe yields difficulty in language comprehension, while sparing the production of grammatical speech. As we shall see, findings from neuropsychology support the existence of relatively autonomous spheres of human intellectual functioning.

The second source of evidence is the study of special populations such as *idiot savants*, prodigies, and autistic children. These individuals exhibit a highly uneven profile of abilities, thereby allowing us to examine an intelligence in relatively pure form. As the work of Feldman (1980, 1986a) has shown, although prodigies may progress through one or more domains with rapid speed, their development in other domains is typically unexceptional. Neither traditional IQ theory nor Piagetian theory can adequately account for the existence of prodigies.

Another criterion is whether an intelligence has one or more basic information-processing mechanisms which operate on various kinds of input. These core operations can be defined as neural mechanisms or computational systems which are genetically programmed to be activated by particular kinds of information. For example, linguistic intelligence encompasses syntactic, semantic, and pragmatic capacities, while musical intelligence includes both rhythmic and pitch abilities.

A fourth criterion is the existence of a distinctive developmental history, through which all normal and gifted individuals pass. The developmental trajectory of each intelligence is relatively independent of the others. So although language develops to a fairly high degree of competence in almost all human beings, artistic skill may require more nurturance to reach an equivalent level. Moreover, the rates of development also vary. Thus, prodigies before the age of ten are more commonly found in some domains, such as music or mathematics, than in others, such as language or art (Feldman, 1986a, 1986b). Each intelligence should also be linked to an identifiable set of "endstates," or expert

performances which represent valued roles in the society. For example, end-states in music include composer, instrumentalist, conductor, and critic.

Fifth, Gardner examined the plausibility of an evolutionary history for each intelligence. A candidate intelligence gained credibility if its evolutionary antecedents could be identified, including whether there existed capacities which could be shared with other organisms. For example, tool use in higher primates, an antecedent of bodily and mechanical skill, dates back several million years (McGrew, 1977). With regard to music, recent studies indicate that parallels can be drawn between the mechanisms underlying development of human musical capacities and the development of birdsong (e.g., Nottebohm, 1980).

A sixth kind of evidence comes from the results of studies of transfer and generalization of skills. Many experimental tests show the difficulty of establishing transfer of learning across tasks (see Pressley, Snyder, & Cariglia-Bull, 1987; Scribner & Cole, 1981; Thorndike & Woodworth, 1901; but cf. Perkins & Salomon, 1989). And psychological processes such as memory and perception may well turn out to be specific to various kinds of input (see, e.g., Fodor, 1983; see also Kosslyn & Koenig, 1992; Warrington, 1986).

A seventh source comes from psychometric findings. Theoretically, tasks that are designed to assess one type of intelligence or its core operations should correlate highly with one another. Of course, the interpretation of psychometric findings is not always straightforward. Many tests call for the use of linguistic and logical-mathematical skills in the solution of the test items, regardless of the type of ability they are designed to measure. In addition, many tasks can be solved using means other than those they are intended to tap. For example, identifying rotations of a figure can be solved by logical-mathematical or linguistic as well as by spatial reasoning.

Finally, an intelligence must also be susceptible to encoding in a symbol system—a culturally contrived system of meaning which captures important forms of information. All humans communicate and construct meaning through the use of such symbolic forms as words, gestures, numbers, and musical patterns. Examples of symbol systems related to the different intelligences include language, mathematics, and picturing.

Based on the above criteria, Gardner identified seven discrete intelligences. The term *intelligence* can be used in two senses: as a capacity which we all have, for example, each one of us has the seven intelligences to varying degrees; and as an honorific, for example, some of us may exhibit high musical or spatial intelligence. These meanings should not be conflated. It is important to keep in mind that, except for unusual populations such as those mentioned above (e.g., *idiot savants*), intelligences never operate in isolation. In any adult role of significance, one is always looking at a combination of intelligences. For example, skilled lawyers draw upon linguistic, logical-mathemat-

ical, and interpersonal intelligences, while carpenters use a combination of spatial, logical-mathematical, and bodily skills.

Before describing each of the intelligences, it may also be useful to distinguish our use of the term *intelligence* from the concepts of *domain* and *field*. In his 1983 book, Gardner did not sufficiently differentiate the three concepts, a lapse which led to unnecessary confusion. Thanks to our continuing collaboration with David Feldman and Mihalyi Csikszentmihalyi, we can now make a systematic distinction (Csikszentmihalyi, 1987; Csikszentmihalyi & Robinson, 1986; Feldman, 1986a). *Domain* refers to the structure of an area of knowledge or human accomplishment, such as mathematics, chess, or literature. These bodies of knowledge develop and change over time. For example, advances in technology may fundamentally alter the nature of a domain such as neuroscience or artificial intelligence. Some domains are well-defined and map relatively directly onto an intelligence, such as mathematics or music; others like politics or the visual arts are not as easy to define and do not lend themselves as well to analysis and evaluation in terms of specific intelligences.

The characteristics of a domain also change over the course of one's development. Thus, in music, talented adolescents must make the transition from an intuitive understanding of musical properties to a more formal knowledge if they are to advance in the domain (Bamberger, 1982). This formal analysis may draw on linguistic or logical-mathematical as well as musical intelligences.

The *field* refers to the social organization of a domain—the roles, behaviors, institutions, and standards related to a domain in a particular society. For instance, the field of music includes the subdivisions of composing, performing, conducting, musicology, and music education. Professional associations, competitions, conferences, journals, awards, and training institutes also make up part of the social organization of the music field. A further component is the people who influence the structure of the domain. Musicians in our culture are influenced not only by other musicians, teachers, and critics, but by orchestra managers, recording producers, public and private funding sources, and the media.

The relation among intelligences, domain, and field can perhaps be best conveyed through an example. Because of our genetic potential to realize “logical-mathematical intelligence,” individuals all over the world explore the numerosity of objects. Many cultures have developed a formal discipline which encompasses the study of number and number-related topics; in the West, this *domain* includes arithmetic, algebra, calculus, topology, and the like. Individuals with high potential in the logical-mathematical area are often drawn to these topics, which are encountered in school or through the study of texts. Determination of who is most accomplished in the domain of mathematics, decisions about what constitutes breakthroughs in knowledge, and even the

determination of who should make judgments about quality are all the concern of the *field* of mathematics.

Thus, while all of the intelligences build on biopsychological potential, they can only be realized in specific tasks embedded in the domains and fields of a culture. We therefore include in our introduction of each intelligence examples of the endstates which depend upon a high degree of a given intelligence, such as poets in language or composers in music. We now briefly describe each intelligence in turn.

Linguistic Intelligence

Linguistic intelligence is the most universally shared and most thoroughly studied of all the intelligences. By the age of 4 or 5, all normal children learn to speak with relative fluency. The core operations of language include sensitivity to the meaning of words (semantics), the order among words (syntax), the sounds and rhythms of words (phonology), and the different functions of words (pragmatics). While syntactic and phonological processes are more strictly “linguistic” operations, semantics and pragmatics likely include inputs from the other intelligences (e.g., logical-mathematical and interpersonal). The autonomy of linguistic intelligence is most clearly demonstrated by the specific localization of the brain mechanisms underlying language in the left hemisphere. Moreover, one can specify lesions in the brain that cause particular difficulties in phonological discrimination, in the pragmatic uses of speech, and in the semantic and syntactic aspects of language. Linguistic intelligence takes different forms in different cultures, ranging from the Koranic scholar to the literary critic. The set of endstates in our culture includes poet, novelist, lawyer, and historian, among others.

Musical Intelligence

Musical intelligence develops earliest of all the intelligences—the talents of musical prodigies are often apparent by the age of 2 or 3. The core operations of music include sensitivity to pitch, rhythm, and timbre. Studies of the brain indicate that the processes and mechanisms underlying music are localized in most normal individuals in the right hemisphere. Different cultures may emphasize one component over another, for example, rhythm is central in sub-Saharan Africa, whereas pitch proves dominant in some Asian cultures. Because music is not as valued in our culture as linguistic and logical-mathematical intelligence, there is little nurturance of musical ability after formal schooling begins, except for children with unusual talent or opportunities. However, in cultures such as China, Japan, and Hungary, musical competence is highly prized: Children are expected to develop proficiency in singing and,

if possible, in instrumental performance as well. In Japan, the Suzuki method is responsible for training large numbers of children to play musical instruments to a fairly high degree of mastery. As noted above, musical intelligence is evident in such endstates as composers, instrumentalists, singers, and sound engineers. Birdsong suggests an apparent link to other species.

Logical-Mathematical Intelligence

Unlike music and language, logical-mathematical ability is not linked to the auditory/oral system. Although many people have speculated about the possible link between musical and logical-mathematical ability, the core operations of each intelligence are fundamentally different. The development of logical-mathematical intelligence has been most thoroughly studied by Piaget (1983). According to Piaget, the basis for all logical-mathematical thought inheres initially in the manipulation of physical objects. Eventually, these actions become internalized and can be conducted inside one's head. In the final stages of intellectual development, the child is able to operate not only on objects or mental images of objects, but on words and symbols as well. Such symbol-manipulating capacities are critical to higher branches of mathematics, where symbols can represent objects, relations, functions, or other operations.

The domains of numbers, mathematics, logic, and science, although not equivalent, do seem to entail related competences. However, while mathematicians are interested in exploring abstract systems for their own sake, scientists are more interested in coming up with ways to model and to explain physical reality. In contrast to language and music, we know comparatively little about the evolutionary antecedents of mathematical ability and its organization in the brain. Endstates in our culture include physicist, chemist, computer programmer, and accountant.

Spatial Intelligence

Both logical-mathematical and spatial intelligences are "object-based" competences. Core capacities of spatial intelligence include the ability to perceive forms and objects, the ability to manipulate or transform the objects or forms, and the ability to produce a graphic likeness of one's visual experience, even without reference to the original stimuli. Thurstone (1938) was one of a number of psychometricians who argued for the existence of a separate spatial ability. Indeed, this is the one intelligence, apart from linguistic and logical-mathematical, which most students of intelligence testing concede should be considered a discrete form of intellect.

The processing of spatial information involves the posterior portions of the right hemisphere of the brain. Lesions to the right parietal regions cause diffi-

culties in visual attention, spatial representation and orientation, imagery production, and memory. Right hemisphere patients who try to use linguistic strategies to aid themselves rarely succeed. Few child prodigies are found in this intelligence, although there are occasional *idiot savants*, such as the autistic child artist, Nadia. In terms of evolutionary history, spatial skills were primary to the group life of many primates for such purposes as hunting, gathering, and traversing wide spaces. Endstates for this intelligence include architect, navigator, sculptor, chess player, painter, and carpenter.

Bodily-Kinesthetic Intelligence

This is the third form of intelligence connected to the use of objects. The two core components of bodily intelligence consist of the ability to use one's body for expressive or goal-oriented purposes, and the capacity to work skillfully with objects, through either fine or gross motor movements. Skilled use of one's body has been important in the history of the species for thousands of years. Higher primates have been using simple tools for several million years. Although in the West the physical and mental realms are often divided, with the body considered inferior to the mind, other cultures do not make such distinctions. In Bali, where all individuals learn to focus attention on bodily features, a sense of balance and fine motor control develop to a high level in almost everyone. In our society, dancers, athletes, actors, artisans, and inventors exhibit aspects of this intelligence.

With regard to neurological evidence, the left hemisphere controls motor activity in most individuals. Injuries to selected areas of the left hemisphere can cause apraxias of great specificity. Apraxic individuals are cognitively and physically capable of carrying out a sequence of actions, but are nonetheless unable to do so. On the other hand, we find neuropsychological patients whose linguistic and/or logical capacities have been severely impaired, but who are capable of performing highly skilled motor activities.

The Personal Intelligences

Gardner also proposes the existence of two personal intelligences. The core capacity of *interpersonal intelligence* is the ability to understand other individuals, particularly their moods, temperaments, motivations, and intentions. The ability to work with and relate to others, whether in the capacity of leader, negotiator, facilitator, caregiver, or friend, is central to this intelligence. We find examples of interpersonal intelligence in politicians, religious leaders, therapists, teachers, and actresses.

Intrapersonal intelligence involves having a veridical picture of oneself, including one's strengths, weaknesses, hopes, and desires. It includes the

capacity to operate adaptively based on such self-knowledge. It can also refer to the ability to draw upon one's affects or emotions as a means of understanding and guiding one's behavior. We see instances of this intelligence in people who are highly knowledgeable about their likes, dislikes, and intellectual profile, and who are able to use this knowledge in directing their life course. Intrapersonal intelligence is also evident in artists, writers, and other individuals who retain ready access to their inner lives and are able to reflect on and express their understanding in meaningful and insightful ways.

With regard to neurological research, the personal intelligences appear to be localized in the frontal lobes. Although our knowledge about the localization of the personal intelligences in the brain is less extensive than our understanding of some of the more standardly computational intelligences, defects in the frontal lobes can interfere with the development of inter- and intrapersonal knowledge. Autistic children seem to suffer from a deficient interpersonal intelligence, while children with Down's syndrome are surprisingly capable of establishing effective relationships relative to their impaired linguistic and logical abilities.

Studies from other cultures present highly contrasting views of the "self." For example, in Bali, virtually all aspects of personal existence are stylized: Individuals are often conceived of in terms of the masks they wear in ritual activities. While the Balinese emphasize the public side of the self, in Morocco the private side of the individual is strictly guarded. Activities such as religion and marriage are carefully insulated from the public realm. While the two personal intelligences are clearly linked in many ways, each one has an identifiable core, a typical pattern of development and breakdown, a range of identifiable endstates, and a characteristic neurological representation.

These, then, are the seven proposed intelligences. The list is not intended to be definitive, but rather to reflect the range of human cognitive potentials as well as the range of roles and products which have come to be valued in our various cultures. The development of the intelligences in different individuals always reflects a combination of environmental influences and hereditary factors. Although the intelligences are rooted in biological potential, to understand individual competence fully, we also need to take into account the cultural context in which the individual is working. It is to this more contextualized aspect that we now turn.

CONTEXTUALIZING THE THEORY OF MULTIPLE INTELLIGENCES

Recently, scholars in the cognitive sciences have proposed that intelligence or cognition is better described as "contextualized" or "distributed," rather than as an entity that operates in isolation in the individual, irrespective of the locale

in which he or she lives (A. Brown, 1990; J. S. Brown, Collins, & Duguid, 1989; Pea, 1990; Perkins, 1990). Let us first outline the various ways in which intelligence can be considered *contextualized* (or *situated*). Learning and knowledge cannot be separated from the activity and culture in which they take place (Brown, Collins, & Duguid, 1989; Rogoff & Lave, 1984). For example, intelligences can be situated in a variety of contexts ranging from home, school, and the workplace, to the wider community and nation. One would not expect to find the spatial and logical-mathematical abilities of a chess master in a culture which did not include chess among its cultural offerings. Moreover, even if chess were part of the general culture, but was not included in a child's home or community, the chances of a child revealing an intellectual proclivity for chess would be slim.

Another way in which intelligence can be considered contextualized is to view it as the product of an interaction between the biological proclivities of individuals, on the one hand, and the needs, values, and opportunities provided by societies, on the other. As we saw earlier, different cultures emphasize different sets of intelligences. Thus, in traditional societies, which depend on cooperative efforts to secure such basic needs as food and shelter, intelligence is associated with the ability to maintain social ties. However, in modern literate societies, intelligence is associated more with linguistic and logical abilities. The manifestation of intelligence can perhaps best be thought of as a dynamic between individual competences and the social structures and institutions that do (or do not) support the development of those competences (Kornhaber, Krechevsky, & Gardner, 1991).

A third way in which intelligence is contextualized is via the evaluative judgments formed by a culture or subgroup regarding both the areas of knowledge and the quality of performance in a particular domain (Goodnow, 1990). For example, mathematics in our society is often considered the province of males rather than females. This is conveyed in both subtle and not-so-subtle ways—from the young schoolgirl who tries to hide her mathematical aptitude in order to remain popular with her peers to those of us who unconsciously direct our questions calling for quantitative knowledge to men.

Related to the idea of intelligence as contextualized is the notion of intelligence as distributed. In the latter view, intelligence is seen not as residing only in the head of the individual, but rather as distributed among other people, resources, and artifacts (Pea, 1990; Perkins, 1990). When individuals attempt to solve a problem of some complexity, we characteristically turn to cultural tools and aids of the domain—materials, strategies, and techniques—which play a critical role in carrying out the task. Except perhaps in school, individuals rarely engage in meaningful problem solving in isolation from one another (Resnick, 1987). For instance, in figuring out a budget, one might use a calculator or computerized spread sheet, look at past budgets, confer with other knowledgeable individuals, and so on. Indeed, our skills to some extent can be

considered specific to the tools which we have inherited (Rogoff, 1990). To take one example, Japanese abacus experts use internalized representations of an abacus to perform complicated mathematical calculations (Hatano, 1982; Stigler, Barclay, & Aiello, 1982).

The relationship between contextualized and distributed intelligence is intimate, but the two concepts are not identical. In speaking of intelligence as contextualized, we highlight the fact that all intellectual performances are an interaction of cognitive potentials and societal needs and opportunities (cf. Olson, 1970). In speaking of intelligence as distributed, we stress that nearly all intellectual performances involve the use of physical and cultural materials that exist in the community—and that these artifacts are part of the distribution of intelligence. In theory, one could have theories of intelligence which are contextualized, but not distributed; or distributed, but not contextualized; in practice, however, contextualized and distributed views tend to co-occur.

How might we then go about envisaging MI theory in terms of a contextualized framework? In traditional intelligence testing, the emphasis falls on solving isolated problems or providing short answers to predetermined questions. This might be adequate preparation for testing in the school context, but it is of limited relevance outside of school. In the workplace and in our personal lives, the projects we undertake are often tied to particular, situation-specific goals, which involve collaborative effort over time. Thus, one way to contextualize the intelligences is to look at how they are deployed across a range of settings, including school, work, home, and extracurricular activities. Rather than examining intelligence in the abstract, we need to examine how the combinations of intelligences are used in carrying out tasks in a variety of circumstances.

A theory of multiple intelligences in context also implies that not all intelligences are drawn upon equally in all settings. For example, the importance placed on the social intelligences differs, depending on the context. Thus, for example, interpersonal intelligence plays a more important role in the public workplace than in the case of the poet in her chamber. But even occupations or vocations which are primarily solitary, such as writing poetry or composing, require interpersonal savvy in seeking funding, working with an agent or manager, or establishing contacts in the field—and of course, the actual works that are produced depend on the creator's knowledge of the world of other human beings. Certainly, most aspects of human life involve interactions with other people, whether family, friends, co-workers, or neighbors.

Another aspect of context concerns the extent to which an individual's own views about self are considered of consequence within a particular culture. In our culture, the ability to understand one's own strengths and limitations is regarded as perhaps of equal importance as the ability to work with others. For example, the awareness that one learns better one-on-one than in a group, or that diagrams and models facilitate understanding more than prose, is useful in

almost any learning situation. A related capacity is the ability to reflect on one's progress, and to use that information in planning subsequent steps, or in rethinking and revising past ones. As we shall see, one way schools can foster the development of intrapersonal intelligence is to help children discover their particular intellectual profile in a broad range of areas, rather than just linguistic and logical-mathematical. In a culture which values original work, intrapersonal intelligence is particularly crucial. However, in a culture which is primarily mimetic, or where the role of self is deliberately downplayed, intrapersonal intelligence will be less valued or will have to be exercised with caution.

A theory of contextualized intelligences also underscores the importance of considering which intelligences and roles should be emphasized, given a particular culture's priorities. If school is designed to prepare students for life upon leaving school, then certain implications follow. For example, although it is important for schools to recognize the various kinds of intelligence which each one of us exhibits, and to give all children a chance to succeed, we also have to consider the context of the field. If only a limited number of slots exist for aspiring singers, dancers, or painters, then it is questionable whether students should be encouraged to pursue related career paths. And if linguistic and mathematical competence do in fact represent important skills for many aspects of adult life, then a relative emphasis on those intelligences might not be misplaced. Thus, considerations of the field must be taken into account when deciding which intelligences will be of primary concern in an educational system.

INTELLIGENCES IN SPECIFIC CONTEXTS

The theory we have presented has a number of implications for different contexts including, but not limited to, the school setting. In any use of the concept of intelligence, a context is always either stated or assumed. Some cultures consider slowness and care to be a sign of intelligence, whereas cultures like our own typically associate intelligence with speed. Beginning with Binet's intelligence tests, the context for intelligence in the West has been school-related tasks and knowledge. However, although academic intelligence may be of critical importance to those who are charged with devising such instruments (i.e., psychologists and other research scientists), it does not have the same importance across the range of worthwhile endeavors (Neisser, 1976). Thus, for example, if businesspeople, religious leaders, heads of households, or artists had created the first intelligence tests, they would have devised very different kinds of instruments—if they would have favored instruments at all.

Perhaps reasons still exist to create a test which will better predict successful performance in school. But we think that a sufficient number of such tests

have already been invented, and it does not pay to devote limited resources toward designing a test to predict academic success 1% more effectively. What we do *not* have are ways of thinking about the uses of the mind across different contexts. MI theory suggests a radically different notion of assessment that has important implications, given Boring's (1923) old saw that "intelligence is simply what the tests of intelligence test." What would it mean to assess intelligences that are contextualized and distributed?

Human activities outside the school setting are, by definition, already highly contextualized: Most involve a combination of intelligences, interaction with other individuals, timely feedback, and the use of tools and other physical aids. They are also tied to specific needs and goals, which are often remote from scholastic concerns. Although traditional theories of intelligence may be related to certain kinds of school knowledge, they are even less relevant to intelligences in operation at the workplace, on the street, or in the home. MI theory provides a framework which should allow us to analyze each one of these settings in its own terms.

For example, in many work settings involving more than one person, it is potentially of great use to be able to identify the intellectual profile of each person as well as of the group as a whole. Tasks could then be delegated according to each person's strengths and weaknesses. Performance would be evaluated not on one, but along a number of dimensions, reflecting the various types of intelligence called on for the particular job. The group profile could be surveyed to determine whether other forms of expertise were needed. The range of intelligences could also be drawn upon for determining the most effective training strategies. Some employees might learn or work better through spatial or experiential means, whereas others might prefer verbal explanations. Finally, as noted earlier, certain intelligences such as the interpersonal are likely to be highlighted in the workplace. Getting along with one's supervisor or staff, with customers, clients, and colleagues can all be critical to effective job performance. No matter how skilled a person is at working alone, at some point one is likely to have to depend on others for successful completion of a task or product.

Cultural values also influence what is considered superior job performance. Certain forms of thinking or approaches to problem solving are judged more elegant or productive than others. As Goodnow (1990) points out, among Anglo-Saxons, ideas which are expressed in a cool and rational fashion are considered more elegant than discursive and emotional forms of expression. In this country, maximized efficiency, individual competition, and material reward dominate the work ethic, whereas in Japan achievement in the workplace is linked to identification with one's firm and a sense of personal responsibility to society.

In such nonwork contexts as family and recreation, the personal intelligences will also assume great importance. Family members are expected to be

responsive to each other's needs, an important component of interpersonal intelligence. A new parent must be sensitive to the nonverbal signals of the infant, while the person who lives alone benefits from an intrapersonal awareness of her own lifestyle preferences. In team sports, players must work together to capitalize on the strengths of each individual. No single person is expected to embody all of the qualities which make up a successful team. Thus, although bodily-kinesthetic intelligence may be at a premium in a sport such as basketball, interpersonal and logical-mathematical competence may also be called into play. There are also a number of physical, mental, and human tools which aid effective performance: video replays, strategy plans, coaches, managers, team members, and other vehicles of distributed intelligence.

MI in the School Context

In our own laboratories, we have explored various options for transporting a contextualized and distributed view of intelligence to the school setting (see, e.g., Gardner, 1993; see also Kosslyn & Koenig, 1992). In this regard, our work can be compared with other efforts to assess scholastic intelligence. However, in contrast to previous efforts, we favor a far more expanded view of performances of significance in school. In particular, we are interested in schools which develop a range of human competencies, which give students numerous ways in which to display their strengths, which foster growth in domains and disciplines valued by the culture, and which develop abilities that can be of use not only in school, but in other contexts of society. In light of these considerations, any treatment of intelligence in the context of school ought to encompass four issues, which we discuss below with reference to some current examples of alternative assessment and curriculum efforts.

What opportunities currently exist for engaging the various intelligences?

As we saw in the opening scenarios, some schools are already attempting to broaden the areas of knowledge to which children are exposed. However, particular consideration needs to be given to whether some intelligences, or aspects of intelligences, merit greater attention than others. Project Spectrum, a collaboration between Harvard Project Zero and Tufts University, is a curriculum and assessment project that was designed to elicit a broad range of capabilities in preschool children (see Krechevsky, 1991; Krechevsky & Gardner, 1990a; Malkus, Feldman, & Gardner, 1988, for details). Spectrum developed 15 different assessments in the areas of math, language, science, music, movement, art, and social understanding. These measures, ranging from structured tasks to observational checklists, are administered in the child's classroom over the course of the school year. The collection of information over time in the child's own environment is a deliberate attempt to blur the line between curriculum and assessment. At the end of the year, all of the information on each

child is written up in a "Spectrum Profile." The report describes the child's intellectual profile, focusing on strengths, with recommendations for home, school, and wider community activities.

In a comparison between results on the Spectrum activities and the Stanford-Binet Intelligence Scale (4th ed.), the Spectrum battery yielded more jagged profiles of the children (see Krechevsky & Gardner, 1990a, for a full discussion). The Stanford-Binet scale did not predict successful or unsuccessful performances either across or on a subset of Spectrum activities. While these data are based on a small sample size ($n = 17$), and the results should be considered tentative, they do suggest that Spectrum engages more areas of intelligence than traditional measures. The findings might also discourage teachers from making broad generalizations about a child's intelligence or problem-solving ability.

Are there opportunities for children to become involved in extended, authentic inquiry, rather than learning information which is relevant only to one-shot, short-answer tests (Perkins, 1990)? According to our theory, children's learning will be better assessed if they are engaged in rich projects over time, ones which are tied to meaningful, real-world activities. Another way of phrasing the same idea is to link scholastic tasks to meaningful endstates in a culture. In Spectrum, we soon realized that intelligences could not be measured in the abstract. Instead, we provided children with a rich environment, including a wide range of engaging materials such as musical instruments, mechanical gadgets, and storytelling boards. The materials were selected to reflect various adult endstates. For example, the assembly task relates to the role of mechanic, the musical instruments to the role of music performer, and the storyboards to the role of novelist or dramatist. Teachers are able to assess unobtrusively their students' particular strengths and interests as they arise in the context of meaningful activity, and then make informed curriculum decisions based on their observations.

Another way in which Spectrum contextualizes its assessments is to use measures which are "intelligence-fair." Rather than viewing all abilities through the window of language and logic, like most standardized tests, "intelligence-fair" measures tap abilities directly, via their own particular medium. Thus, children sing and play musical instruments in the music assessment, and participate in creative and athletic movement sequences in the assessment of bodily-kinesthetic intelligence.

In addition to intellectual competence, working styles such as a child's persistence, attention to detail, and level of confidence are also addressed. The contextualized MI framework prompts teachers to consider different areas of competence to determine whether a child's working style may be domain-specific, perhaps surfacing only in an area of strength.

Is there ample opportunity for collaborative work? Children's understanding grows as a result of interactions with adults and other children. The work

of Brown and Palincsar (1989) on the teaching of reciprocal reading, and Lampert (1986) and Schoenfeld (1985) on mathematical problem solving, among others, provide good examples of the power of collaborative learning and group interaction. These projects incorporate a variety of collaborative learning techniques, such as small-group discussion, brainstorming, and peer critique. The potential benefits of collaborative learning are many: it emphasizes the active and constructive nature of learning; it provokes discussion and articulation of positions; it encourages the use of metacognitive skills; and it helps students appreciate the various roles and processes involved in complex problem solving (see Collins, Brown, & Newman, 1989, for a full discussion of successful collaborative learning projects). This kind of group exchange is also critical to many work situations. Group problem solving fosters the awareness that we all have different intellectual profiles, and that a person's contributions cannot be judged along a single dimension. A contextualized MI theory provides a framework for identifying alternative ways to explore problems via the seven intelligences.

Is there sufficient scaffolding for children's growth in a domain, in particular, with regard to distributed cognition? Children need help in knowing how to draw upon other "distributed" resources, whether they are human—teachers, librarians, specialists—or material—books, calculators, written documents, or computer databases. It is helpful in this regard to consider Greeno's (1990) definition of a subject-matter domain as a collection of resources for the cognitive activities of knowing, understanding, and reasoning, rather than a structure of facts, concepts, and principles. The Practical Intelligence for School Project (PIFS), a collaboration between Harvard Project Zero and Yale University, represents an attempt to help children in middle school understand and manage the distributed aspects of intelligence. The goal of the PIFS Project is to teach students strategies for negotiating the complex demands of school (see Gardner, Krechevsky, Sternberg, & Okagaki, 1993; Krechevsky & Gardner, 1990b; Sternberg, Okagaki, & Jackson, 1989; Walters, Blythe, & White, in press, for fuller discussion).

Many of the PIFS units focus on the distributed nature of intelligence by explicitly addressing the topic of human and material resources. These units encourage students to consider the types of resources available for writing a paper or working on a math problem, the reliability of using different resources, their particular advantages and disadvantages, and so on. Comparisons are drawn between resources which we use in our daily lives (books, television, recipes, maps) and those needed for a particular school subject, like math. Students are also asked to reflect on their personal experiences and interests outside of school, and to consider similarities and differences with their scholastic experiences. The units also foster such practical skills as revising a paper, choosing a project, taking notes, and organizing and presenting one's work. In line with the notion of distributed intelligence, PIFS emphasizes

the importance of drawing on human, personal, and physical resources as an important part of problem solving, both in and outside of school.

CONCLUDING NOTE

In conclusion, we attempt to situate MI theory both with respect to other theories of intelligence and to other applications of intelligence theory. Standard intelligence theory has emerged out of psychometrics, with its attendant trait-centered view of cognition. Many students of intelligence think of intelligence as composed of different factors, which they then try to isolate through administering various tests and applying correlational procedures. MI theory stretches in the direction of three other disciplines: neurobiology—the brain sources of cognition; cognitive science—identification of modules and their constituent processes; and cultural studies—ethnographic sensitivity to the abilities valued elsewhere. MI theory also attempts to capture new conceptualizations of cognition, such as the modularity of specific cognitive faculties and the notion of intelligence as distributed, as well as insights from the study of different cultures and contexts. Overall, we intend our theory to be an expanding and unifying conception, rather than one which directly confronts or refutes psychological trait and factor analytic approaches.

As for the practical side, in contrast to most traditional conceptions and assessments of intelligence, which may bear only limited relevance to curriculum planning and design, Gardner's theory has turned out to be applicable to educational practice. The theory also provides a framework for analyzing the range of roles and skills valued in a society and determining whether some intelligences are being neglected (for instance, most cultures would likely profit from a balanced emphasis on intra- and interpersonal intelligences). Third, the theory enables us to consider whether individual intellectual profiles in a society match the roles and functions important to that society, with implications for such policy-level determinations as allocation of resources and establishment of educational priorities. Finally, the MI framework pluralizes the endpoints of development, increasing the chances that each person will have his or her own profile of abilities recognized and valued.

REFERENCES

- Bamberger, J. (1982). Growing up prodigies: The mid-life crisis. *New Directions for Child Development*, 17, 61–78.
- Boring, E. G. (1923). Intelligence as the tests test it. *New Republic*, 35, 35–37.
- Brown, A. L. (1990, April). *Distributed expertise in the classroom*. Paper presented at the American Educational Research Association, Boston, MA.

- Brown, A. L., & Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), *Knowing, learning and instruction: Essays in honor of Robert Glaser*. Hillsdale, NJ: Erlbaum.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Leadership*, 18(1), 32–42.
- Ceci, S. J. (1990). *On intelligence... more or less: A bio-ecological treatise on intellectual development*. Englewood Cliffs, NJ: Prentice-Hall.
- Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser*. Hillsdale, NJ: Erlbaum.
- Csikszentmihalyi, M. (1987). Society, culture, and person: A systems theory of creativity. In R. J. Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives*. Cambridge, UK: Cambridge University Press.
- Csikszentmihalyi, M., & Robinson, R. (1986). Culture, time and the development of talent. In R. Sternberg & J. Davidson (Eds.), *Conceptions of giftedness*. New York: Cambridge University Press.
- Eysenck, H. J. (1981). *The intelligence controversy*. New York: Wiley.
- Feldman, D. H. (1980). *Beyond universals in cognitive development*. Norwood, NJ: Ablex.
- Feldman, D. H. (1986a). *Nature's gambit*. New York: Basic Books.
- Feldman, D. H. (1986b). Giftedness as a developmentalist sees it. In R. Sternberg & J. Davidson (Eds.), *Conceptions of giftedness*. New York: Cambridge University Press.
- Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Gardner, H. (Ed.). (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.
- Gardner, H., Krechevsky, M., Sternberg, R. J., & Okagaki, L. (1993). Intelligence in context: Enhancing students' practical intelligence for school. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice*. Cambridge, MA: Bradford Books.
- Goodnow, J. J. (1990). The socialization of cognition: What's involved? In J. Stigler, R. A. Schweder, & G. Herdt (Eds.), *Cultural psychology: Essays on comparative human development*. Cambridge, UK: Cambridge University Press.
- Gould, S. J. (1981). *The mismeasure of man*. New York: W. W. Norton.
- Greeno, J. G. (1990, February). *Number sense as situated knowing in a conceptual domain* (Rep. No. IRL 90-0014). Palo Alto, CA: Institute for Research in Learning.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Hatano, G. (1982). Cognitive consequences of practice in culture specific procedural skills. *Quarterly Newsletter of the Laboratory of Comparative Human Cognition*, 4, 15–17.
- Jencks, C. (1972). *Inequality*. New York: Basic Books.
- Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, 39, 1–123.

- Jensen, A. R. (1980). *Bias in mental testing*. New York: The Free Press.
- Kosslyn, S., & Koenig, O. (1992). *Wet minds*. New York: Free Press.
- Kornhaber, M., Krechevsky, M., & Gardner, H. (1991). Engaging intelligences. *Educational Psychologist*, 25(3 & 4), 177–199.
- Krechevsky, M. (1991). Project Spectrum: An innovative assessment alternative. *Educational Leadership*, 48(5), 43–48.
- Krechevsky, M., & Gardner, H. (1990a). The emergence and nurturance of multiple intelligences: The Project Spectrum approach. In M. J. A. Howe (Ed.), *Encouraging the development of exceptional skills and talents*. Leicester, UK: The British Psychological Society.
- Krechevsky, M., & Gardner, H. (1990b). Approaching school intelligently: An infusion approach. In D. Kuhn (Ed.), *Developmental perspectives on teaching and learning thinking skills*. Basel: Karger.
- Lampert, M. (1986). Knowing, doing, and teaching multiplication. *Cognition and Instruction*, 3, 305–342.
- Malkus, U., Feldman, D. H., & Gardner, H. (1988). Dimensions of mind in early childhood. In A. Pellegrini (Ed.), *The psychological bases of early education*. Chichester, UK: Wiley.
- McGrew, W. C. (1977). Socialization and object manipulation of wild chimpanzees. In S. Chevalier-Skolnikoff & F. E. Poirier (Eds.), *Primate bio-social development: Biological, social and ecological determinants*. New York: Garland.
- Neisser, U. (1976). General, academic, and artificial intelligence. In L. B. Resnick (Ed.), *The nature of intelligence*. Hillsdale, NJ: Erlbaum.
- Nottebohm, F. (1980). Brain pathways for vocal learning in birds: A review of the first 10 years. *Progress in Psychobiological and Physiological Psychology*, 9, 85–124.
- Okagaki, L., & Sternberg, R. J. (1990). Teaching thinking skills: We're getting the context wrong. In D. Kuhn (Ed.), *Developmental perspectives on teaching and learning thinking skills*. Basel: Karger.
- Olson, D. R. (1970). *Cognitive development*. New York: Academic Press.
- Pea, R. (1990, April). *Distributed intelligence and education*. Paper presented at the American Educational Research Association, Boston, MA.
- Perkins, D. (1990, April). *Person Plus: A distributed view of thinking and learning*. Paper presented at the American Educational Research Association, Boston, MA.
- Perkins, D., & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher*, 18, 16–25.
- Piaget, J. (1983). Piaget's theory. In P. Mussen (Ed.), *The handbook of child psychology, Vol. 1: History, theory and methods*. New York: Wiley.
- Pressley, M., Snyder, B. L., & Cariglia-Bull, T. (1987). How can good strategy use be taught to children? Evaluation of six alternative approaches. In S. M. Cormier & J. D. Hagman (Eds.), *Transfer of learning*. New York: Academic.
- Resnick, L. (1987). Learning in school and out. *Educational Researcher*, 16, 13–20.
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York/Oxford, UK: Oxford University Press.
- Rogoff, B., & Lave, J. (Eds.). (1984). *Everyday cognition: Its development in social context*. Cambridge, MA: Harvard University Press.
- Sattler, J. M. (1988). *Assessment of children* (3rd ed.). San Diego, CA: Sattler.
- Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando, FL: Academic.

- Scribner, S., & Cole, M. (1981). *The psychology of literacy*. Cambridge, MA: Harvard University Press.
- Spearman, C. E. (1923). *The nature of "intelligence" and the principles of cognition*. London: Macmillan.
- Sternberg, R. J., Okagaki, L., & Jackson, A. (1990). Practical intelligence for success in school. *Educational Leadership*, 48(1), 35–39.
- Sternberg, R. J., & Wagner, R. K. (Eds.). (1986). *Practical intelligence: Nature and origins of competence in the everyday world*. Cambridge, UK/New York: Cambridge University Press.
- Stigler, J. W., Barclay, C., & Aiello, P. (1982). Motor and mental abacus skill: A preliminary look at an expert. *Quarterly Newsletter of the Laboratory of Comparative Human Cognition*, 4, 12–14.
- Thorndike, E. L., Bregman, E. O., Cobb, M. V., & Woodyard, E. I. (1926). *The measurement of intelligence*. New York: Teachers College.
- Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review*, 8, pp. 247–261, 384–395, 553–564.
- Thurstone, L. L. (1938). *Primary mental abilities*. Chicago, IL: University of Chicago Press.
- Walters, J., Blythe, T., & White, N. (in press). PIFS: Everyday cognition goes to school. In H. Reese & J. Puckett (Eds.), *Advances in lifespan development*. Hillsdale, NJ: Erlbaum.
- Warrington, E. K. (1986). Memory for facts and memory for events. *British Journal of Clinical Psychology*, 25, 1–12.

Chapter 12

***CAM: A Theoretical
Framework for Cognitive
Abilities Measurement***

Patrick C. Kyllonen

Armstrong Laboratory, Human Resources Directorate

INTRODUCTION

Over the past several years, I and my colleagues in the Armstrong Laboratory's¹ Learning Abilities Measurement Program (LAMP) have been conducting basic research on individual differences in cognition. This chapter reviews some of that work, discusses weaknesses in our current understanding of human abilities, and proposes a new theoretical framework for cognitive abilities measurement. Although the research discussed here is broad, I believe it retains a very simple focus. Quite simply, all our research has been directed towards developing a comprehensive battery of ability tests. On first consideration, this may seem rather atheoretical. One does not need a theoretical framework to assemble a batch of off-the-shelf tests for various purposes, but this is not our intention. Our goal is to articulate principles for what we should measure and how

¹Formerly, the Air Force Human Resources Laboratory.

we should measure it, with a special emphasis on comprehensiveness. The battery is a realization of those principles. In this way, the battery is like a theory. It is available to be inspected, criticized, elaborated upon, rejected, and revised, both in its underlying foundations and in its current realization.

The construction of a principled test battery requires addressing several basic research issues. First, we need to determine what abilities there are. Then, once we have identified an ability, the question is, how general is it? Is the ability something that only comes into play on a few highly specialized tasks, or is it a ubiquitous ability, something that enters into a broad range of tasks? Finally, once we have identified an ability, and determined that it is a fairly general one, the question is, what is the most efficient way to measure it in individuals?

Components of Individual Differences in Cognition

What exactly is an ability? As I define it here, an ability is a kind of component, which is simply a source of individual differences variance.² Components include broad ability components, knowledge components, task-specific performance components, error components—any person (as opposed to task) variable that produces individual-differences variance. Psychologically, components may be defined as *relatively autonomous structures* (e.g., declarative memory) and *mechanisms* (e.g., storage mechanism in short-term memory) that operate together in accomplishing cognitive activities, such as learning, remembering, problem solving, and decision making. The integrity of these components, that is, their suitability for accomplishing cognitive activities, will vary from one person to the next. Some individuals may be relatively strong in some components, weak in others. The environment may also affect the integrity of the components. People may be less capable of certain kinds of cognitive activities (e.g., problem solving, learning) in certain kinds of environments (e.g., speed-stress, weightlessness, the deep sea) due to component degradation.

An important issue concerns the scope of the components. Components can be identified narrowly, as they are in current neuropsychological work (Posner, 1978). Conversely, components can be identified broadly. Spearman's (1904, 1927) two-factor theory of cognition, for example, identified the *g* factor as a broad component involved in virtually any cognitive activity. A productive framework might run somewhere between these two extremes. It would be useful to identify a set of components such that if we could measure the efficiency of each of them, for a given individual, in a given environment, we could

²Throughout the chapter, I use the terms components and factors almost interchangeably. I tend to use components as the more general term. I tend to use factors in the more technical sense as latent variables that appear in factor analysis models, which causally influence performance on particular tasks.

predict reasonably well whether that individual could perform a set of *characteristic tasks*. Characteristic tasks could be defined as those an individual might actually be called upon to perform in some routine classroom, training, or performance setting.

Methodology

A key theoretical issue concerns methodology for identifying cognitive components. It is convenient to separate past investigations into top-down and bottom-up approaches (Table 1). The top-down approach, historically aligned with confirmatory, experimental psychology, and more recently with artificial intelligence investigations, involves rational analyses of performance tasks. The task analyst addresses the question: What components must be posited for such-and-such task to be accomplished? A classic example of this approach is Gagne's (1965) learning taxonomy; a contemporary version is Anderson's (1983) task analysis of problem solving in geometry and computer programming. The advantage to the top-down approach is that components identified through careful task analyses and validated over experimental manipulations are precise and well understood, to the point where simulation programs embodying these components can be developed (e.g., Anderson's ACT* system). The disadvantage is that what is gained in depth is achieved at the cost of breadth. Rational task analyses yield in-depth descriptions of performance components, but on a narrowly defined universe of tasks. Consequently, the field knows a great deal about a small number of tasks—arithmetic, geometry, programming. How the components of these tasks map on to others is a question awaiting considerable research effort.

The bottom-up approach involves exploratory individual-differences investigations. Factor theorists (e.g., Carroll, 1989; Cattell, 1971; Guilford, 1967; Spearman, 1927; Thurstone, 1938) have attempted to define the core factors of a broad range of human cognition through exploratory factor analysis of matri-

Table 1. Comparison of Approaches to Component Identification

	<i>top-down</i>	<i>bottom-up</i>
approach to identifying components	rational, confirmatory	empirical, exploratory
methodology	experimental	correlational
component acceptance criterion	sufficiency	necessity
advantage	depth	breadth
disadvantage	narrow	blind

ces of cognitive test intercorrelations. The advantage of this approach lies in its breadth. Rather than limiting their scope to accounts of performance on a small number of tasks, factor theorists posit broader components to account for variability on a wider variety of intellectual tasks. The chief disadvantages relate to the exploratory nature of these efforts. The constitution of components identified through factor analytic investigations depends to some degree on the tasks for which they are found to account. Different sets of tasks yield different components—such is the nature of bottom-up empiricism. Further, the components that are robustly identified in study after study, such as reasoning ability, are derided for being poorly understood (Johnson-Laird, 1985).

There can be a fruitful synthesis of the top-down and bottom-up approaches. Individual-differences (correlational) methods can be employed to validate components arising from psychological theory associated with experimental approaches. This differs from traditional abilities investigations in that correlational methods are used in primarily a confirmatory fashion. The logic, elucidated by Underwood (1975) is this: If x is a general component of human cognition, then (a) it should be possible to measure x in a context outside that in which it was originally developed, and (b) different measures of x should be correlated over persons. Failing to satisfy either condition casts doubt on the premise.

Top-down and bottom-up approaches provide complementary perspectives on component necessity and sufficiency. Top-down approaches, through detailed modeling of performance on individual tasks, generate lists of components necessary for performance of those tasks. Bottom-up approaches, by attempting to account for performance on a broad range of tasks, conversely address the question of whether the positing of a particular set of components is sufficient, or comprehensive, for accounting for a broad range of cognition.

A Consensus Information-Processing Model

A consensus information-processing model would be an ideal source from which to hypothesize components. Figure 1 displays the outlines of a model that has been employed by a number of cognitive researchers (e.g., Anderson, 1983; Card, Moran, & Newell, 1983; Langley, 1983). This framework has served as the basis for our own research over the last several years. The system includes a short-term *working memory* and two long-term memories, *procedural memory* and *declarative memory*. Information from the environment is interpreted by the *perceptual processing system*, and a record of that information is deposited in working memory. Information in working memory activates related information in both declarative and procedural memory. This is equivalent to saying that information in those long-term memories is retrieved into working memory by the *cognitive processing system*. One type of infor-

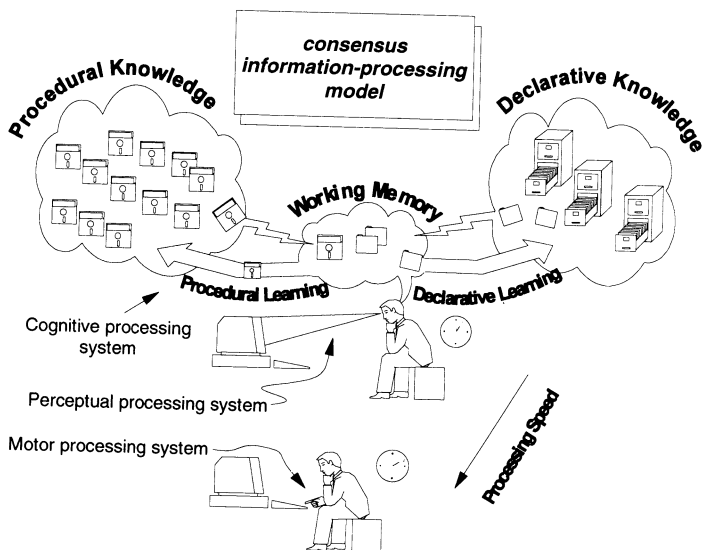


Figure 1. Consensus information processing model. The system includes a short-term, *working memory* and two long-term memories, *procedural memory* (whose contents are depicted as floppy disks) and *declarative memory* (whose contents are depicted as files in file cabinets). Information from the environment (the computer display is interpreted by the *perceptual processing system*, and a record of that information is deposited in working memory.

Information that can be active in working memory is an instruction for the system to take some overt action, such as pressing a response key. The *motor processing system* handles the chore of interpreting the instruction and executing the response.

In generating a model of individual differences from the consensus information-processing model, we need to keep in mind two overarching considerations. First, are the components listed in Figure 1 sufficient for characterizing (or accounting for) processing on a diverse sample of cognitive tasks? Or are there important components or subcomponents missing from the figure? Second, given a list of components, such as those identified in Figure 1, what is the range of individual differences observable on each one? Documenting these differences is required for developing adequate models of task performance. And noting covariances in individual differences over different components in the consensus processing model should shed light on the underlying structures and mechanisms giving rise to those differences.

THE FOUR-SOURCES MODEL

Our initial theoretical framework for abilities measurement, which we called the four-sources model, was based on the idea that the most important (ubiquitous) sources of individual differences on cognitive tasks were the components of the consensus information-processing model (Kyllonen & Christal, 1989). Specifically, we proposed that individuals differ in four basic ways: the capacity of working memory, the execution speed of the various processing systems (cognitive, perceptual, motor), the breadth of declarative knowledge, and the breadth of procedural knowledge.

The four-sources model was developed from a rational analysis of the requirements of an information processing system. We intended that it be used as a starting point for thinking about what kinds of abilities to test for in studies that investigate the correlations between basic human abilities and accomplishment in natural (ecologically valid) training and on-the-job performance situations (cognitive correlates studies). This heuristic function of the four-sources framework has served us well, at least in the sense that it has stimulated the development of literally hundreds of novel ability tests (Kyllonen, in press). We have argued (Kyllonen, 1990; Kyllonen & Christal, 1989) that not only can many of the individual differences factors long recognized by psychometricians be mapped onto the four-sources factors, but the four-source model suggests new factors.

To validate the four-sources framework, with respect to comprehensiveness, my LAMP colleagues and I have conducted a number of studies examining how much of the individual-differences variability on a variety of short- and intermediate-term learning and performance tasks can be accounted for by four-sources measures. We have examined learning logic gates (Christal, 1990; Kyllonen & Stephens, 1990; Kyllonen & Woltz, 1989), computer programming (Pena & Tirre, 1990b; Shute & Kyllonen, 1990), if-then rule learning (Chaiken, 1990; Woltz, 1988), paired-associates learning (Kyllonen, Tirre, & Christal, 1991; Kyllonen & Tirre, 1988; Tirre, in press), and reading comprehension tasks (Tirre, 1990; Woltz, 1990a). On average, we find that we can account for 40% to 50% of the individual-differences variance on such learning tasks. Most encouraging, from an applied perspective, is that the four-source measures typically account for substantially (10% to 30%) more of the between-subjects variance in learning than do conventional tests (i.e., the Armed Services Vocational Aptitude Battery, ASVAB) currently used operationally for personnel selection and classification (Christal, 1989).

This constitutes a validation of sorts for the four-sources model. On the other hand, accounting for 40% to 50% of the variance is a long way from providing a full account of individual differences in learning. Ideally, we would like to account for *all* the individual differences variance on any kind of intellectual task. It may be absurd to think that we can achieve this goal; still, there

are some benefits in identifying a clear-cut target to strive toward. The utility, as a cognitive task analysis system, of any component taxonomy is only as good as its ability to predict how well any random individual will perform on any arbitrary cognitive task. Providing fairly complete accounts of individual differences variance on a diverse set of cognitive tasks is a prerequisite to developing such a system.

PROBLEMS WITH THE FOUR-SOURCES MODEL

The problem with the four-sources framework is that it is too simple. It is predictably wrong that people differ in just four basic ways, and our failure to account for much more than half the variance on various learning tasks supports this prediction. The implication is that we must develop a more elaborate taxonomy of sources of individual differences. It is instructive to review a prediction model from one of our studies (Kyllonen & Stephens, 1990) as a way of identifying specific loci of the inadequacies of the four-sources framework.

Figure 2 shows the path diagram identifying the relationships among four-source and learning variables in a study on acquisition of skill in solving logic gate problems. Following convention, circles denote latent factors; squares denote observed variables; e represents error in the variables; D represents disturbances (error) in the factors; arrows denote the causal direction of the regression: independent variable (source of variance) to dependent variable. Note that we included seven learning variables, all of which we let load on a general Declarative Learning (DL) factor, and three of which we let load additionally on an orthogonal procedural learning (PL) factor. We included two variables to measure each of the four source factors (WM = working memory, PK = procedural knowledge; DK = declarative knowledge; the processing speed factor was dropped from the analysis). The depicted model fit the actual data (variance-covariance matrix of all 13 measures) fairly well as indicated by the standard goodness-of-fit criteria $\chi^2 = 10.4$, $p = .312$.

The four-source factors together accounted for 56% of the variance in the DL factor ($1 - D^2$), and (coincidentally) 56% of the variance in the PL factor. In both cases, almost all of the prediction accuracy was due to the WM factor. While intriguing in its own right (for a discussion, see Kyllonen & Woltz, 1989), this is not directly pertinent to the issue at hand. This indicates that there is a considerable amount of uniqueness (44% of the variance) in the DL and PL factors, independent of any reliability problems with the particular tests we used as four-source measures and learning indicators. That is, learning abilities on the logic gates task are related to but also somewhat distinct from non-learning cognitive abilities (WM, PK, DK, and PS). An obvious question is this: How much more of the variance in the DL and PL factors, which are specific to the logic gates domain, would be accounted for if we had included a

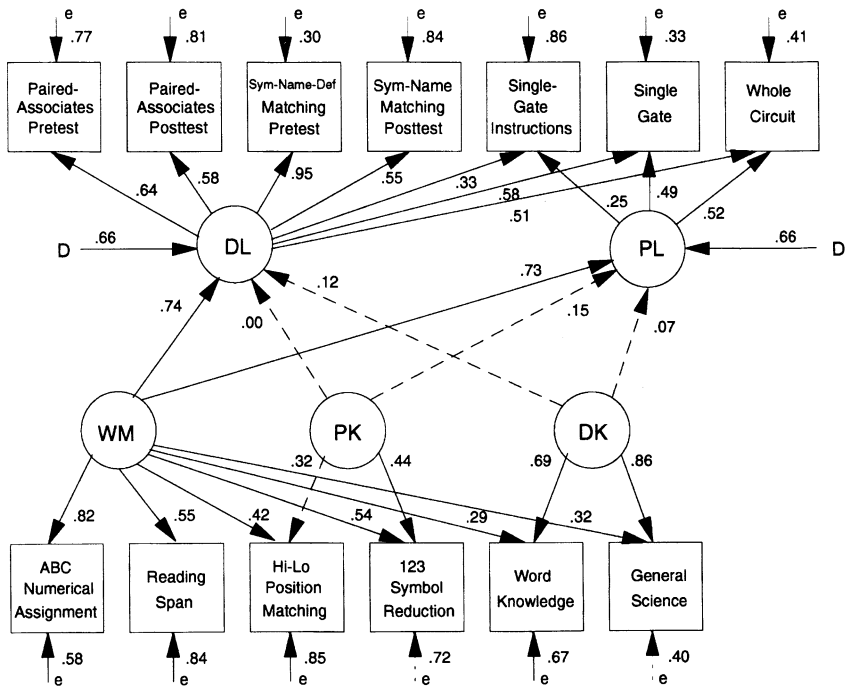


Figure 2. Path diagram identifying the relationships between aptitude and learning variables from the Kyllonen–Stephens study on acquisition of skill in solving logic gate problems. DL = Declarative Learning; PL = Procedural Learning; WM = Working Memory capacity; PK = Procedural Knowledge; DK = Declarative Knowledge. Circles = latent factors; squares = observed variables; *e* = error in the variables; *D* = disturbances (error) in the factors; arrows = the causal direction of the regression: independent variable (source of variance) to dependent variable.

general DL and general PL factor in the model? (The DL and PL factors in the model are specific to learning logic gates.) Presumably, considerably more. Thus, an obvious improvement to the four-sources model would be the inclusion of additional *general-DL* and *general-PL* factors. At the very least, inclusion of such factors would enable us to determine how much of the individual-differences variance in learning a particular task is predictable from individual differences in general learning ability (as opposed to domain-specific learning abilities).

Another feature of the figure is that it indicates two (in some cases three) sources of variance for every test: error and factor (in some cases two factors). Error constitutes anywhere between 9% (Symbol-Name-Definition Matching Pretest) and 93% (Single-Gate Instructions) of the variance, averaging roughly

50%. Although not indicated on the figure, this 50% “error variance” can be further subdivided into test unreliability and test uniqueness (bias or method variance; Humphreys, 1976). In addition to factor uniqueness (discussed in the immediately preceding paragraph), these two sources of test uniqueness represent additional inadequacies in the four-source model. Ideally, we would like to account for a considerable chunk of the 50% error variance in terms of a priori, specifiable factors.

Need More Components

In addition to general learning factors there may be other general factors whose inclusion would similarly improve the four-sources model. Carroll (1990) has identified 16 factors in an analysis of the ETS kit of cognitive references tests, a widely used battery in research settings. Some of those factors may be mapped to four-source factors. For example, Raymond Christal and I have demonstrated that working memory capacity and general fluid reasoning ability are largely indistinguishable (Kyllonen & Christal, 1990). But many of the conventional factors (e.g., closure speed) may not map onto the four-source framework (Kyllonen, 1990). Still other factors, besides those identified by Carroll, have been identified in the performance literature, such as *temporal processing* (e.g., dynamic spatial ability; Hunt, Pellegrino, Frick, & Alderton, 1988) and *time-sharing* (or multitasking, or information coordination; Brookings, 1990; Pellegrino, Hunt, & Yee, 1989) ability. An obvious way to improve the four-sources model, with respect to comprehensiveness, is by appending these other factors to the four-sources list.

Need More Subcomponents

Positing additional components to account for more criterion task variance is only one approach to increasing variance accounted for. A complementary approach, which has motivated much of the newer work on information processing conceptions of intelligence, is based on the idea that criterion and predictor tasks already are samples from the same task space. The reason for lack of perfect prediction is that every task is a different mixture of more basic subcomponents. For example, suppose that in a given study there is a predictor task measuring working memory and there is a criterion task that also reflects working memory. Suppose that the two tasks are not perfectly correlated after controlling for measurement error. One explanation is that the criterion task reflects additional components. This explanation motivates a search for those additional components. Another explanation is that the two tasks vary in their mixture of subcomponents. For example, if working memory capacity is a composite of more basic, separable subcomponents (e.g., momentary capacity, decay, and

resistance to interference), an individual's scores on two working memory tests could differ depending on how the two tasks reflected the subcomponents. The degree of difference between the two tasks would depend on the degree to which the subcomponents are separable (independent), and the degree to which the tasks varied in the importance of the different subcomponents.

Content (Domain) Distinctions

Researchers conducting exploratory factor analyses of multiple aptitude test batteries (e.g., Thurstone, 1938; Wothke et al., 1990) have typically identified at least some factors dominated by the content feature of tests (e.g., the verbal factor, the quantitative reasoning factor), but have downplayed or overlooked the possibility that pure content factors per se could be identified. A pure content factor would be a source of variance independent of, rather than confounded with a process source (e.g., working memory, processing speed). Exploratory factor analysts have typically identified confounded process-content factors (e.g., quantitative reasoning). Psychologically, a pure content factor would reflect the possibility that people differ in relative knowledge (e.g., verbal vs. quantitative) independently of general knowledge differences, and that these relative knowledge differences affect performance on all kinds of processing tasks (memory, decision making, reasoning, etc.).

Reanalyses of multiple aptitude batteries using methods other than exploratory factor analysis can identify content sources of variance. For example, a multidimensional scaling reanalysis of Thurstone's (1938) primary mental abilities matrix reveals separate process and content dimensions along which tests can be arrayed (see Figure 3; adapted from Snow, Kyllonen, & Marshalek, 1984). I have repeatedly found that confirmatory factor-analytic models of the interrelations among cognitive tests are improved with the inclusion of domain factors (e.g., verbal, spatial, quantitative). For example, consider Figure 4, which shows the results for a model of the interrelationships between Reasoning (R and RA), Working Memory (WM), General Knowledge (GK), and Processing Speed (PS) factors (from Kyllonen & Christal, 1990). The model was improved significantly when verbal and quantitative content factors, orthogonal to the factors depicted, were added. For example, verbal analogies and 3-term series, both verbal content tests, were allowed to load additionally on a verbal content factor (not shown). Augmenting the model with content factors yielded a significantly better fit to the data.

Input Modality Distinctions

According to the consensus information-processing model (Figure 1) individuals *perceive* information from the environment, then *encode* a representation of

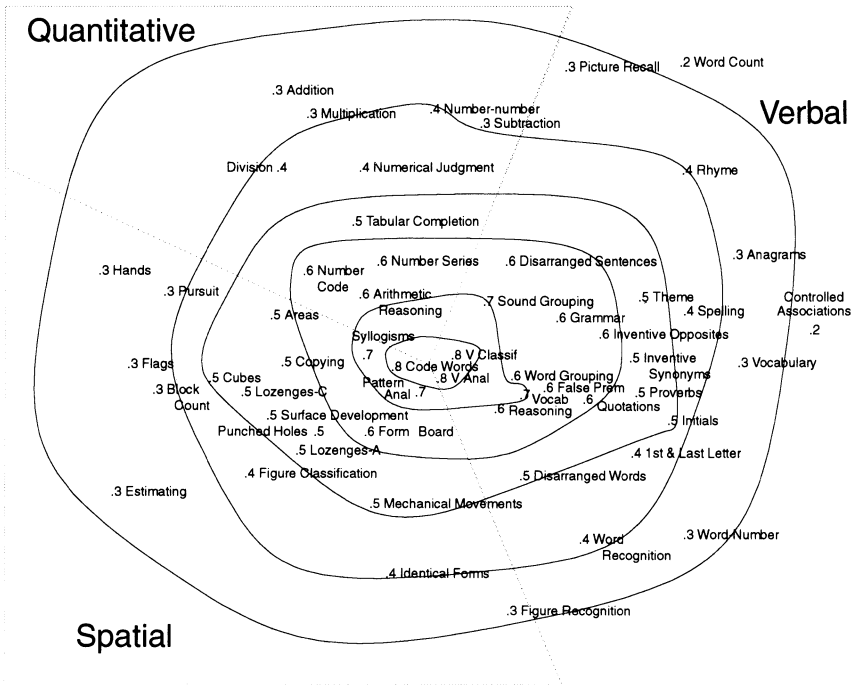


Figure 3. Multidimensional scaling reanalysis of Thurstone’s (1938) primary mental abilities correlation matrix (adapted from Snow, Kyllonen, & Marshalek, 1984). Coefficients next to test names are first-factor loadings from an unrotated principal axis factor analysis of the matrix. Lines encircle variables with same-magnitude loadings. Wedges (Verbal, Spatial, Quantitative) are based on inspection of tests’ contents.

that information in working memory. The four-sources model treats the perceptual encoding process as a given, ignoring any variance in test performance attributable to the perception process itself. This is clearly an expedient simplification. It is likely that people vary in how adeptly they encode information (i.e., how quickly they form a working memory representation, and how rich that representation is), and such differences are likely to vary by modality. That is, skillful visual encoders may be poor auditory encoders, and vice versa.

Thus far we have not collected data on how input modality affects processing, partly because we expect that this variance source is not as important as more purely cognitive sources are. But another reason is primarily technological: Until recently, we simply did not have voice recognition capabilities on our systems (we now do). Modality will probably not make any difference for some processes. For example, examinees are unlikely to be affected differen-

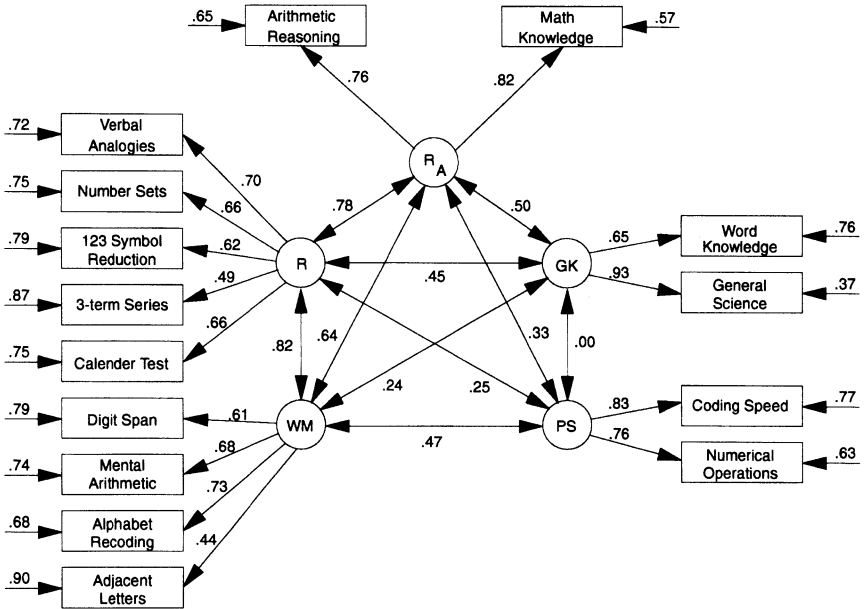


Figure 4. Confirmatory factor-analysis model of the interrelationships between Reasoning (R and RA), Working Memory (WM), General Knowledge (GK), and Processing Speed (PS) factors (adapted from Kyllonen & Christal, 1990).

tially by whether they are asked “What is the name of the process by which plants make food from sunlight?” (a declarative knowledge item) by the computer display screen or through headsets. For other tasks, such as working memory capacity tasks, input modality may affect examinees differentially. Reading and listening comprehension scores are highly correlated ($r = .7$ to $.9$; Gernsbacher, Varner, & Faust, 1990; Palmer, MacLeod, Hunt, & Davidson, 1985; Sticht, 1972), but not so highly as to preclude the identification of variance attributable to modality.

Response Format Distinctions

For most kinds of tasks of interest to us, variance produced by individual differences in motor responding skill is undesirable. In these cases, we seek response formats that minimize individual differences. Touch-screen responses are direct and natural for examinees (albeit tiring), and requiring them may eliminate nuisance variance due to subjects occasionally forgetting the arbitrary response mappings that must be assigned when using keyboards (Detterman,

personal communication, June, 1990). Vocal responding is another minimal variance response format: For many tasks, issuing voice commands (true, false, three, high) is probably more automated than is pressing keys on a keyboard. This would result in fewer “inadvertent” errors, and consequently “cleaner” scores. The research question is this: For what tasks are scores contaminated by the existence of inadvertent errors induced by response format? A useful study would identify the kinds of cognitive tasks for which response format (voice vs. touch-screen vs. keyboard) affects scores, and properties of those scores such as reliabilities and validities.

In some cases, response variance is theoretically interesting. For example, psychomotor abilities have historically played an important role in air crew selection research (Bordelon & Kantor, 1986; Carretta, 1989). Part of psychomotor ability is purely cognitive (i.e., response-format independent), but another part is undoubtedly response-specific. The four-sources model ignores motor response variance as well as input modality variance. Yet, here again, there may be response components, and there may be differential response components, such that some performers are relatively adept at discrete motor responses (e.g., pressing keys), whereas others are relatively adept at continuous motor responses, and still others are relatively adept at voice responses. Most generally, response format is a potential source of variance independent of the process, domain, and modality sources thus far discussed. It is an open question how many response format sources there may be (e.g., continuous vs. discrete, fine vs. gross).

Component Interactions

Interaction factors are sources of individual differences independent of the component factors that make up the interaction. For example, an individual may be relatively adept (skillful compared to others) at performing continuous tracking tasks (e.g., tracking a moving object on the display screen with a mouse), and relatively adept at a spatial working memory task, yet be poor at performing the two tasks simultaneously, compared to other individuals. In fact, there is a suggestion in the literature that coordination of multiple information sources and response requirements is an ability fairly independent of one’s ability to perform the tasks independently (Brookings, 1990; Pellegrino, Hunt, & Yee, 1989; cf. Ackerman, Schneider, & Wickens, 1984). Any of the components thus far discussed could theoretically interact with any other component (e.g., working memory \times processing speed; knowledge \times input modality). An interaction exists to the extent that scores on tasks performed in isolation do not predict scores on a task that combines the two (or more) component tasks. Because many real-world tasks are essentially multiple component tasks, it is important to document the circumstances under which such interactions emerge.

Table 2. Types of Component Interactions

<i>forms of interaction</i>	<i>locus of effect</i>	
	<i>task</i>	<i>individual differences</i>
superadditive	WM × PS	WM × TPs
subadditive	WMs × WMv	WMv × GKv

It is useful to organize the kinds of interactions that could occur between tasks in the form of a 2×2 table (Table 2). There are two possible forms an interaction may take. A *superadditive* interaction is one in which the two components have an effect greater than the sum of their individual effects; a *subadditive* interaction is one in which the two components have an effect less than the sum of their individual effects. The effect could be observed as changes in *mean performance on a task*, or it could be observed as an orthogonal *individual differences factor* (i.e., where interaction task performance was not predictable from an additive combination of component task performances). Cell entries in Table 2 are the kinds of components that we predict might lead to such interaction effects.

For example, consider a superadditive interaction realized on a task (Working memory [WM] × Processing Speed [PS]). Suppose the probability of error (Pe) on some working memory task were .3, and Pe on some processing speed task were .1. An error on a combined task could be a result of either an error on the WM task, the PS task, or both, and thus by simple probability, the expected Pe (combined) = $1 - (.7 * .9) = .37$. A superadditive interaction would be demonstrated if Pe (combined) > .37. This could happen, if for example, the PS task requirements constituted a significant enough increase in working memory load to “push people over the threshold” of what they could temporarily store, in a “chaos-type” process. In the same way, combining a dynamic spatial task (“space invaders” type task) with a working memory task has been shown to produce a superadditive interaction; but in addition, individual differences in performance of this kind of combined task have been shown to be independent of performance on the component tasks (Pellegrino et al., 1989).

Subadditive interactions are probably rarer, but they could result from a situation in which (a) performance on a task is not impaired much (if at all) in the presence of having to perform another task; and (b) performance on one task actually facilitates performance on the other, in some sense. For example, two working memory tasks performed simultaneously, one with spatial stimuli, one with verbal stimuli, might not interfere with one another, and might even be mildly facilitative if the stimuli can be integrated somehow. An individual differences factor might emerge if some were more skillful at the integrative process, for example.

Test Scores and the Speed–Accuracy Problem

In four-source investigations, we typically measure *accuracy* (e.g., percent correct) on tests of working memory, declarative knowledge, and procedural knowledge, and we measure response *latency* on tests of processing speed. Yet subjects can almost capriciously raise their accuracy scores by simply choosing to spend more time on problems, or lower their latency scores by choosing to respond less accurately. Subjects' ability estimates in these investigations are thus biased to the degree that subjects vary in where they position themselves on their own, personal speed–accuracy tradeoff functions. For example, suppose subjects cleanly divided into “button pushers” and “compulsives.” Button pushers would consistently score poorly on tests and compulsives would consistently score well. We would thus find high intercorrelations among tests, and conclude there was a strong general ability factor. If so, we would have committed the mistake of attributing performance commonality to ability when it should have been attributed to speed–accuracy tradeoff (i.e., a button-pushing versus compulsive style). More generally, if positioning is stable across tasks we end up with an inflated general factor estimate (i.e., inflated relative to what it would be if everyone positioned themselves at the same accuracy or latency point on the speed–accuracy tradeoff); if positioning is stable across task groups (e.g., working memory tasks) we estimate an inflated group factor, if positioning changes from test to test, we overestimate the uniqueness component of a test; and so on.

Ideally we would like a single proficiency score on an individual for a test. This involves holding either accuracy or latency constant, then treating the free variable as the score. In conventional testing (paper-and-pencil), this is accomplished by scoring number correct per unit time (e.g., 5 minutes). In computerized testing, there are more possibilities (see Kyllonen, 1991). We can fix limits (deadlines) at the item or the test level, we can model tradeoff statistically (because item response time information is available), we can vary deadlines and statistically model the effect on accuracy over the various deadlines, or we can use payoffs and feedback interactively to get subjects to perform at a target accuracy or latency level.

In addition to the single proficiency score, it may be useful to determine where subjects choose to position themselves when left unconstrained. We simply do not know whether such positioning is habitual (general across all tests) or local to particular types of tests, local to particular tests, or completely haphazard. It may be that some individuals are generally careless (sacrifice accuracy to go fast), others are generally careful (sacrifice speed to be accurate); it may be that these patterns change over time (e.g., day-to-day); or it could be that task characteristics govern these patterns in either general or idiosyncratic ways.

Modeling Error of Measurement: The Causes of Unreliability

Error variance arises from the choice of a particular set of items and from moment-to-moment fluctuations in performance. It has no special theoretical interest. (Cronbach, 1970, p. 320)

A fairly exhaustive list of possible sources of test score variance is provided in Table 3 (adapted from Thorndike, 1947). Sources are either *general* (across all items that could have been developed for this particular test) or *specific* (to the particular test administered); and either *lasting* (affecting performance equally today vs. next week) or *temporary* (affecting performance only today). A table similar to this one could be developed to identify sources of factor variance as well as test variance. For example, a *specific/lasting* source of variance in a factor, such as working memory, is the particular set of tests we develop to measure that factor.

Generally, we interpret a test score as an indicator of ability, but as Table 3 shows, there are numerous other influences on that test score. The challenge is to identify these sources, determine the magnitude of their effect, then develop procedures for either eliminating their influence (if indeed they are of no theoretical interest), or reliably identifying and measuring them in different contexts. For example, the effects of guessing are minimized with many alternative multiple-choice tests, free recall tests, or by lengthening the test. The effects of test wiseness (or general sophistication in understanding test instructions) are minimized by making test instructions extremely simple and straightforward. General and specific affective sources (e.g., general confidence in test-taking ability; a good feeling about this particular test) might not be easily eliminated, but might be identified through a study habits survey (e.g., Dixit, 1991), or through manipulations of goals and feedback (e.g., Kanfer & Ackerman, 1989, 1990).

General vs. specific (item) effects. Internal consistency reliability indicates the effect of the particular items administered. Form-to-form variability is due to the idiosyncracies of items. In a recent study (from data collected in Shute, in press) we administered 30 four-sources tests, on two occasions, separated by 4 to 5 days. This allowed the computation of two reliability estimates: an internal consistency estimate, and a test-retest estimate. Averaging over the tests, internal consistency reliability (for percentage correct scores) was about .8, indicating that about 36% of the variance was item-specific. Interestingly, internal consistency reliability for latency was much higher (average around .95). That is, very little of the latency variance was due to item idiosyncracies—most of the variance was due to general subject characteristics.

It may be possible to predict item idiosyncrasy, to some degree. An account of item effects would essentially be a theory of what makes a good item. Currently, test development relies on empirical trial and error. We design a set

Table 3. Sources of Test Score Variance

	<i>General</i>	<i>Specific (to the set of items)</i>
Lasting	Abilities Affect	Abilities Affect
Temporary	Test-wiseness Health, fatigue, emotional strain motivation/rapport practice heat/light/ventilation	fatigue/ motivation practice

Note. Adapted from *Essentials of Psychological Testing* (p. 175) by L.J. Cronbach, 1970, New York: Harper & Row. Reprinted by permission.

of items, according to a test definition, then collect data on those items, then discard items with poor psychometric properties, specifically, low item total correlations, restricted variance, extremely high or low mean performance levels, and so on. In numerous studies, we have informally investigated good versus bad items as determined by item-total correlation data. For example, we recently analyzed a general knowledge survey and a working memory test. From casual inspection, two item features stood out as correlates of item badness. Many of the bad items were either close to the floor (too hard), or close to the ceiling (too easy). Other bad items had an oddness property, where a fix to the item is almost immediately suggested. A challenge is to develop a priori principles for good items, thereby reducing the cost of the test-revise cycle, and at the same time expanding the capability for conducting cognitive task analysis.

A parallel analysis can be applied on tests at the factor level. What features make a particular test a good or poor test of working memory capacity? What are the features associated with test uniqueness? We find that some working memory tests consistently load highly on a working memory factor, others do not. Besides average difficulty (as with items, easy and hard tests are more likely to have high uniqueness), a test feature that seems to be associated with uniqueness is the degree to which the operations in the test are “natural,” that is, operations similar to those people perform in their normal training, academic, or workplace environments. An example of a test with high uniqueness is the silly syllogisms test from the ETS kit (e.g., all horses grow in forests, a peach is a horse, therefore all peaches grow in forests, true or false?). The premises given in this test are absurd, and conflict with everyday knowledge.

Stable vs. temporary (occasion) effects. Test-retest reliability indicates the effect of occasion. In the study discussed above, we found a rather large occasion effect: On average, test-retest reliabilities were about 20% lower than

internal consistency reliabilities. Latency is particularly sensitive to occasion—the discrepancy between the two reliabilities was greater for latency than for accuracy. Some of the variance resulting from retesting is due to the same sources as test-retest consistency variance—that is, forms are not the same, and thus may measure slightly different abilities. (If the same forms are readministered, then there might be learning effects.) According to Table 3, effects unique to occasion are (a) change of environment (e.g., retesting is done in another room with better lighting); and (b) the general phenomenon of subjects having good and bad days. Because we held environment fairly constant, the good versus bad day source is worth pursuing.

Test scores for a subject on a good day are slightly inflated; test scores for a subject on a bad day are depressed. The question is, can we determine good versus bad days for a subject independently of their score (if we cannot the explanation is tautological)? A possibility that we have not explored systematically is that mood, as measured by standard surveys (Watson, Clark, & Tellegen, 1988), is a valid indicator of the affective variables that influence the day-to-day fluctuations in test scores. There also is research suggesting that personality interacts with time-of-day and caffeine use to influence test scores by up to .5 standard deviation (Revelle, 1989).

Need Explicit Causal Connections

The four-sources model allows for correlation among the components, but does not specify causal directionality in intercomponent relationships. Nor is the four-sources model specific in how multiple factors can contribute to variance in a particular test. A more complete cognitive abilities model is not simply a more exhaustive list of variance components; rather, it is additionally a specification of how those components are causally connected, and of how any particular test can be analyzed with respect to multiple sources. The specifics of causal models organizing component interrelationships and variance sources in tests are touched upon in the second half of this chapter. However, there are several issues that can be discussed in the abstract.

A distinction that can be made in causal modeling is between indicator and composite factors (cf., Cohen, Cohen, Teresi, March, & Valez, 1990). Both kinds of factors are identified by their tests, but the relationship between the tests and the factors differs. Indicator factors, which are independent variables, determine (causally govern) test performance; composite factors, which are dependent variables, are determined by test scores. Tests associated with indicator factors must be intercorrelated; if they are not, the factor does not have much explanatory power. Tests associated with composite factors do not have to be intercorrelated, and they can even be negatively correlated. An idea that we have been considering is that it may be theoretically productive to mix indicator and composite factors in abilities models to represent which factors are

fundamental, basic level and which are derived. Basic level components would appear as indicator factors; derived components would appear as composite factors. For example, working memory capacity might be treated as a composite factor, with subcomponents of working memory capacity, such as central executive capacity, or verbal articulatory loop capacity, treated as basic, indicator factors. An idea worth pursuing is that an appropriate distinction between indicator and composite factors can “explain” test and factor uniqueness.

Another issue is hierarchical modeling. There is a naive dichotomy sometimes expressed (albeit more likely in face-to-face encounters than in print) between general ability and special abilities models. The four-sources model is seen as a special abilities model, which invalidates it *prima facie* in light of the evidence for a general ability. But this is a false dichotomy, and it is easy to show how mixed models that posit both general and special abilities are not only possible but desirable (Gustafsson, 1989; see also Kyllonen & Christal, 1990, appendix). A general ability can be assumed to underlie or determine performance on all tests administered in a battery; orthogonal special abilities can be assumed simultaneously to determine performance on subsets of tests.

Absolute Versus Relative Measurement

Absolute measurement of ability components is the missing puzzle piece in a theory of abilities. The four-sources model as well as any other contemporary abilities “theory” specifies variance sources, and allows for comparing individuals to other individuals with respect to those sources. But extant theories fall short of specifying in any absolute sense what an individual can do. An absolute measurement system would provide component (factor) scores tied to absolute capabilities, such as “can hold three items in working memory,” “takes 200 ms to retrieve an item from long-term memory,” “has a working vocabulary of 20,000 words” and the like. What current systems provide instead are statements such as “is at the 55th percentile in reasoning ability” or “is two standard deviations below the mean in verbal ability.” Competency testing provides absolute measurement, but is atheoretical, and disconnected from the abilities literature.

At the risk of inviting philosophic argument, I believe that an inability to go beyond relative measurement precludes the development of a true *theory* of abilities, as opposed to merely a taxonomic list of variance sources. The relative-measurement limitation also thwarts progress in applying abilities theory as the basis for a task analysis system. To be useful for designing instruction, task analysis must identify knowledge and skill requirements in an absolute sense (e.g., “can use an oscilloscope”). An abilities theory based on absolute measurement would (ideally) be capable of specifying both (a) the knowledge and skill requirements of a task in an absolute sense, and (b) the knowledge and skill possessed by an individual in an absolute sense. From this informa-

tion, individualized training plans could be developed or, at a minimum, required training time, given task requirements and individual capabilities, could be estimated.

The establishment of an absolute measurement system would incorporate individual-differences research into mainstream cognitive science. In a sense, the advancement to absolute measurement relies on advances in cognitive theory. For example, once we know that learning is characterized by power law improvement, we can estimate and interpret individual differences in absolute rate of improvement (e.g., Woltz & Shute, 1991). In this sense, a true theory of intelligence will be a superset of a general theory of cognition (or some aspect of cognition, such as learning). This is not to say that intelligence research is forever relegated to a derivative role. For, in another sense, individual differences research can inform cognitive theory by identifying components, that is, by specifying what “constants” in human performance we ought to be measuring in the first place.³

The goal of absolute measurement of ability components, like the goal of accounting for all the variance in test performance, seems out of reach at present. Still, we can gauge progress by how much we approximate that goal. An absolute measurement orientation can influence one’s approach to *identifying* components. For example, although working memory capacity and reasoning ability are similar factors (Kyllonen & Christal, 1990), working memory capacity is the more useful construct by the criterion of amenability to absolute measurement. An absolute measurement orientation can also influence one’s approach to *investigating* components. One concerned with absolute measurement seeks psychological explanations for why a facet manipulation of a task (e.g., an increase in study–test delay) results in an average response time increase of, say, 700 ms. To one concerned only with relative measurement, such differences reflect a diverse item pool, with respect to difficulty, but are of no theoretical consequence. With an absolute measurement orientation, explaining task-to-task variability is on par with explaining person-to-person variability. This leads to an attitude shift—one constantly seeks to tie any task facet manipulation to an ability component rather than simply use it as a convenient way to alter difficulty level.

Summary

Although the four-sources model has proved useful in motivating the development of tests, and in accounting for more of the variance in various learning criteria than can be accounted for by some conventional aptitude batteries, the model is fairly sketchy and far from complete. The model would be improved

³I put constants in quotes to signify the stance that these would be individual difference variables rather than constants, although the average value of such variables might be treated as constants in human engineering or other applications.

by positing additional components and subcomponents, and by considering input modality, response format, and content features of tests. It might also be that components interact with one another in subadditive (compensatory) and superadditive (amplificatory) fashion. Test (and factor) scores may be made more reliable by controlling speed–accuracy tradeoff, and, more generally, by systematically identifying variance sources that are presently labeled error.

One utility test for a model of ability components is whether it can serve as the basis for a cognitive task analysis system. We would like to be able to inspect a cognitive task, and identify a priori the sources of variance that influence individuals' ability to perform on it. This capability requires that we first identify potentially important variance components (i.e., sources of variance), particularly, knowledge and skill categories, and develop causal models of how those components influence scores. But in addition to specifying components, a task analysis system must ultimately specify training requirements. To do this, an ability components model must, at a minimum, be capable of specifying task requirements and ability component proficiencies on an absolute scale.

THE CAM TAXONOMY

So far I have discussed the theoretical framework that guided the LAMP project's earlier research on individual differences in cognition, and I have discussed some of the limitations of that framework. Some, but not all, of these limitations are resolved in our current framework, which we call the Cognitive Abilities Measurement or CAM framework, which supersedes the four-sources framework.

Figure 5 presents a proposed four-dimensional taxonomy of cognitive abilities, which we can refer to as the CAM Taxonomy. The vertical dimension lists cognitive processing factors. These include the four-source factors, plus declarative and procedural learning factors, temporal processing, time sharing, a set of unspecified interaction factors, and a set of unspecified other factors. The horizontal dimension specifies three broad knowledge domains—verbal, quantitative, and spatial—which are completely crossed with the cognitive processing factors. Auditory versus visual input (stimulus) modality is specified in the orthogonal third dimension; and response type (keyboard, mouse, voice) is a fourth orthogonal dimension. The cognitive process and domain dimensions of the taxonomy are suggested by past LAMP work; the other dimensions are adapted from Wickens's (1980, 1984) attentional taxonomy.

CAM Taxonomy versus the SOI

The three-dimensional taxonomic diagram invites comparisons with Guilford's (1967, 1985) Structure-Of-Intellect (SOI) model. Recall that Guilford orga-

nized tests in a *content* (visual, auditory, symbolic, semantic, behavioral) by *operations* (cognition, memory, divergent production, convergent production, evaluation) by *product* (units, classes, relations, systems, transformations, implications) taxonomy. The CAM *knowledge domain* is similar to the SOI *content* dimension, except that we omit a behavioral category, replace quantitative for symbolic, and unconfound domain and modality. CAM's *cognitive process* is similar in spirit to the SOI *operations* dimension, but the cells are quite different. The CAM taxonomy has nothing analogous to the product dimension.

The CAM and SOI proposals are similar in that both can be used to specify tests. That is, given a description of a particular taxonomic cell, one should

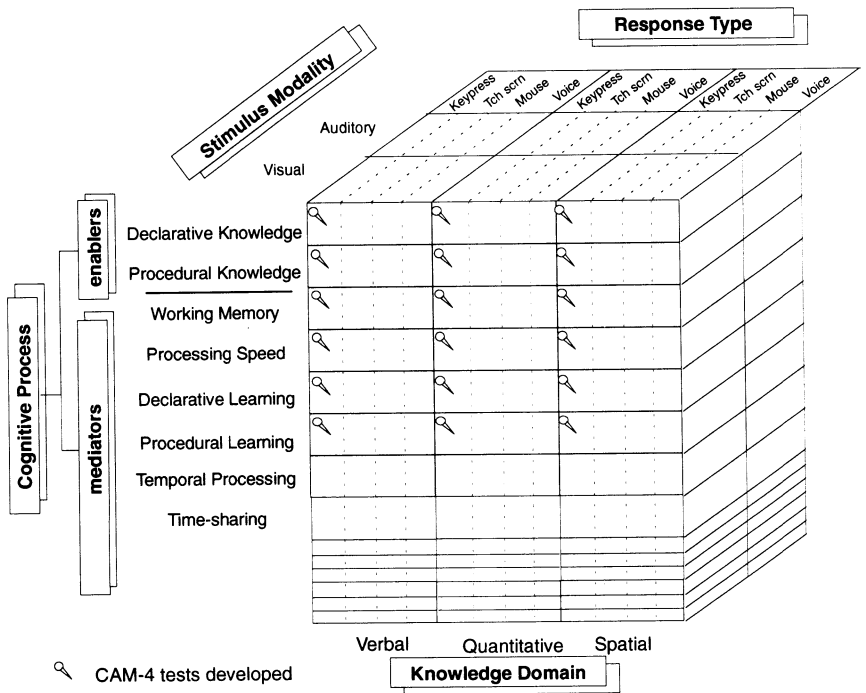


Figure 5. Proposed 4-dimensional taxonomy of cognitive abilities (the CAM Taxonomy). The vertical dimension specifies Cognitive Processing factors; The horizontal dimension specifies three broad Knowledge Domain; The third dimension is Stimulus Modality (Auditory versus Visual Input); and the fourth is Response Type (keyboard, mouse, voice). The cognitive process and domain dimensions of the taxonomy are suggested by past LAMP work; the other dimensions are adapted from Wickens' attentional taxonomy (1980; 1984).

be able to develop a test for that cell. For example, in the SOI, combining cognition operations, semantic content, and a relations product gives a verbal analogies test. In the CAM taxonomy, a test involving working memory, in the verbal domain, with visual stimuli, and a keypress response could also be a verbal analogies test. In both systems numerous other tests could be developed to fit those specifications. The key point is that under both systems, tests can be developed from the specifications contained within the taxonomy itself (cell definitions for the CAM taxonomy are discussed throughout the remainder of the chapter). Conversely, any existing test can be located within the taxonomy by noting its cognitive processing requirements, the domain from which test stimuli are drawn, and so on.

The two proposals differ in the specifics. Still, given that the SOI project is regarded in some circles as a modest failure (Cronbach, 1970; Horn, & Knapp, 1973; McNemar, 1964), a question is why we propose a model that is formulated at least in the same spirit. One response is that access to large numbers of subjects, and a wide community of investigators afforded by Air Force projects (of which the SOI project is an instance) compels this kind of systems-level (some may prefer “grandiose”) thinking. More seriously, let us consider criticisms of the SOI, which cluster into four groups:

1. SOI cells do not map onto constructs used in cognitive psychology.
2. The procrustean statistical methods used to validate the SOI are flexible enough to confirm any hypothesis (Horn & Knapp, 1973), and are therefore inappropriate for theory confirmation.
3. Abilities are organized hierarchically, rather than in accord with a facet model (Cronbach & Snow, 1977).
4. The SOI makes unimportant distinctions.

The first criticism is deserved, but at the same time, the CAM taxonomy is immune to this criticism. The CAM taxonomy better reflects current thinking in cognitive psychology insofar as the cognitive process distinctions either are derived from standard information processing models (the 4 sources and the 2 learning factors) or have emerged in individual differences studies (temporal processing and time sharing). In contrast, many of the distinctions in the SOI, such as those in the Products classification, appear to be rooted in a logical rather than psychological analysis. Empirical analysis confirms this suspicion (e.g., Cronbach & Snow, 1977, p. 157). This criticism of the SOI is not new, but on the other hand, there has been no serious attempt to update the SOI in accordance with developments in cognitive psychology (Guilford, 1982).

The second criticism is specific to a particular method of factor analysis (procrustean) that is no longer viable. But viable, statistically sound validation methods now exist due to developments in confirmatory factor analysis since the time of Guilford's original work. Regarding the third criticism, confirmato-

ry factor analysis is also simultaneously capable of reflecting (and testing the validity of) a hierarchical organization of abilities and a facet structure (e.g., Gustafsson, 1989; see also Kyllonen & Christal, 1990, Study 4), as I discussed previously. Thus, facet and hierarchical abilities models are compatible. It becomes an empirical rather than a methodological issue whether either or both representations are appropriate.

Criticism 4 warrants careful consideration. Cronbach (1970, p. 343) expressed it concisely: "Guilford's distinctions, then, reflect very subtle differences between tasks. The fine subdivision appears to add very little to our ability to explain or predict.... Guilford's is a fine-grain analysis, and fine-grain analyses are not necessarily useful."

Indeed, for many practical purposes, such as personnel selection decisions, it may not be necessary to collect any more data on a person than an estimate of general ability (Ree & Earles, 1990). But the utility of fine-grain distinctions does not rely only on whether additional factors predict more criterion variance. There are other possible applications, including personnel classification ("person-job match"), training treatment assignment or aptitude-treatment interaction studies, cognitive skills training, and performance assessment. And independently of practical issues, it can still be theoretically important to integrate individual-differences work with the main body of cognitive research. One way to attempt this is to represent the constructs of cognitive research in an individual-differences framework, such as is done in the CAM taxonomy.

Evaluation of the CAM Taxonomy

We have now collected data on 59 tests developed to fit into 18 of the CAM taxonomy cells, as indicated by the pins in Figure 5. Table 4 lists the names of these tests. Tests were either adapted from the literature or were developed in-house based on their tractability to taxonomic definitions (see Table 5). Each test consists of two parallel forms, which were administered to approximately 600 subjects on two separate occasions (four days apart). We have not yet fully analyzed these data.

Note in Table 4 that within each cell there are 3 or four test paradigms (e.g., rapid serial classification, verification span). Paradigms are constant across columns. For example, consider the four-term order paradigm (Figure 6). These tests require the examinee to order two elements according to a set of transformation rules, then to order two other elements according to those rules, then to order the two element sets according to those same rules. The verbal domain version of the test involves verbal operations (linguistic transformation) on sentence stimuli; the spatial version involves spatial operations (movement) on spatial stimuli (Wechsler-like blocks), but both tests are identical at the paradigm level (e.g., the experimental design for both tests is the same, manipulating voice, negation, etc.). This design enables a clean test of domain

Table 4. CAM Taxonomy and CAM Battery Test Names

Factor	Paradigm	Domain		
		Verbal (V)	Quantitative (Q)	Spatial (S)
Processing Speed (PS)	1: Two-term Order 2: Verification 3: X Assignment 4: Single Opposites	PSV1: <i>Furniture-Animals</i> PSV2: <i>Silly Sentences</i> PSV3: <i>Common Category</i> PSV4: <i>Up-Down</i>	PSQ1: (10 .. 90) PSQ2: <i>3-Number-Facts</i> PSQ3: <i>Hard Number-Facts</i> PSQ4: <i>10-minus-n</i>	PSS1: <i>Blocks</i> PSS2: <i>Synthesis +</i> PSS3: <i>Synthesis +</i> PSS4: <i>Opp MatrX Sq</i>
Working Memory Capacity (WM)	1: 4-Term Order 2: <i>Verification Span</i> 3: XYZ Assignment 4: <i>Contnu's Opps</i>	WMV1: <i>Furniture-Animals</i> WMV2: <i>Silly Sent+Words</i> WMV3: <i>Categories</i> WMV4: <i>Up-Down</i>	WMQ1: (10 . .90) WMQ2: <i>3-Num-Facts+Dig</i> WMQ3: <i>Hard Num-Facts</i> WMQ4: <i>10-minus-n</i>	WMS1: <i>Blocks)</i> WMS2: <i>Synth+/Matrix</i> WMS3: <i>Synthesis +</i> WMS4: <i>Opp MatrX Sq</i>
Declarative (Fact) Learning (FL)	1: P'rs-Cued-Recall 2: <i>Block-Old-New</i> 3: Symbol Pairs	FLV1: <i>Word Pairs</i> FLV2: <i>Word Blocks</i> FLV3: <i>Noun-Pair Lookup</i>	FLQ1: <i>2-digit Pairs</i> FLQ2: <i>2-digit Blocks</i> FLQ3: <i>Digit-Pair Lockup</i>	FLS1: <i>Palmer-Fig Pairs</i> FLS2: <i>Bruce-Fig Blocks</i> FLS3: <i>Fig-Pair Lookup</i>
Procedural (Skill) Learning (SL)	1: Rap. Ser. Cl'sfn 2: Reduction 3: If-Then	SLV1: <i>Sub-Verb-Adv</i> SLV2: <i>Future-Past-Pres</i> SLV3: <i>Edible-Living</i>	SLQ1: <i>Hi-Lo Num PatrnS</i> SLQ2: <i>Odd-Even</i> SLQ3: <i>Odd-Big</i>	SLS1: <i>4-Square</i> SLS2: <i>Circles</i> SLS3: <i>Shaded Square</i>
Procedural Knowledge (Induction) (IN)	1: Sets 2: Series 3: Matrices	INV1: <i>Word Sets</i> INV2: <i>Word Series</i> INV3: <i>Word Matrices</i>	INQ1: <i>Number Sets</i> INQ2: <i>Number Series</i> INQ3: <i>Number Matrices</i>	INS1: <i>Figure Sets</i> INS2: <i>Figure Series</i> INS3: <i>Figure Matrices</i>
Declarative (General) Knowledge (GK)	(no paradigm manipulations in this category)	GKV1: <i>Nelson-Narens</i> GKV2: <i>Abstract Facts</i>	GKQ1: <i>Distances</i> GKQ2: <i>Counts & Prob'ities</i> GKQ3: <i>Measurements</i> GKQ4: <i>Dates</i>	GKS1: <i>City Directions</i> GKS2: <i>Len's & Angles</i>

Table 5. CAM Factor Definitions

<i>PROCESS</i>	
Processing Speed:	Speed of retrieving, transforming, or responding to stimuli
Working Memory Capacity:	The ability to process and store information simultaneously
Declarative Learning	Ability to acquire new declarative knowledge (commit a novel fact to memory, then strengthen it with use, and minimize decay over time)
Procedural Learning:	Ability to acquire new rules and skill
Declarative Knowledge:	Depth, breadth, accessibility, and organization of declarative knowledge base
Procedural Knowledge:	Depth, breadth, accessibility, and organization of procedural knowledge base
<i>DOMAIN</i>	
Verbal:	Naturally verbal operations (e.g., linguistic transformations, category and synonymy judgments, part-of-speech classifications) performed on verbal stimuli (e.g., words, sentences)
Quantitative:	Naturally quantitative operations (e.g., arithmetic operations, sign reversals, high-low and odd-even judgments) performed on quantitative stimuli (e.g., digits, numbers)
Spatial:	Naturally spatial operations (e.g., rotation, reflection, physical matching, synthesizing) performed on spatial stimuli (e.g., Bruce figures, Palmer figures, partly filled matrices, Wechsler blocks)

effects because domain is unconfounded with paradigm. Previous individual differences studies of this type confound the two.

The validation of the CAM taxonomy as a whole involves two major steps:

1. Compute proficiency scores for each of the tests (and compute internal consistency and test–retest reliabilities of those scores); and
2. Compare various models of the variance–covariance matrix for those scores.

Scoring

The main issue in scoring concerns how we use accuracy and latency information. The simplest possibility is to compute accuracy scores and ignore latencies for all tests except those in the Processing Speed (PS) category, for which we compute latency scores and ignore accuracies. Another possibility is to

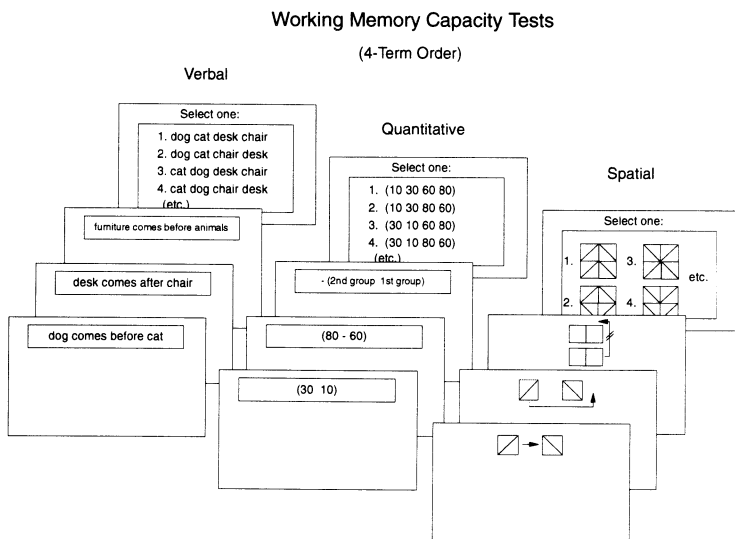


Figure 6. CAM tests based on the 4-term order paradigm. Tests require the examinee to order two elements according to a set of transformation rules, then to order two other elements according to those rules, then to order the two element sets according to those same rules. The verbal-domain version of the test involves verbal operations (linguistic transformations) on sentence stimuli; the spatial version involves spatial operations (movement) on spatial stimuli (Wechsler-like blocks), but both tests are identical at the paradigm level (e.g., the experimental design for both tests is the same, manipulating voice, negation, etc.).

covary out latency (or some transformation, such as log latency) in computing accuracy scores (e.g., percent correct, d' , log errors) and to covary out accuracy in computing latency scores (for the PS tests). Still another possibility is to compute ratio scores, such as accuracy per unit time (e.g., percent correct divided by average response time).

For any of these scores, an issue is what latency score to use in the computations. *Average latency over all items* has the problem of mixing "time to termination" (for incorrect items) and processing time (for correct items). *Latencies on corrects only* has the problem of confounding item-difficulty level with subject (subjects who only get easy items correct will have a low latency because hard items require more processing and consequently take longer). An interesting latency score we have used with some success (Christal, personal communication), which minimizes the effects of the subject \times item-difficulty confound is *item-standardized latency*. This is a person's average standardized latency for correct items, where each standardized latency is with respect to

only those subjects who got that item correct. A variant on this score that removes the confound altogether is *difficulty-adjusted latency*. This method essentially treats latencies on incorrect items as missing values. We compute missing value estimates using other information about subjects and items, employing standard linear modeling techniques (and either least-squares or maximum likelihood estimation procedures).

Estimating a proficiency score is clearly a problem in itself. A particular problem is determining which of these and still other possible proficiency scores is best. An ideal validation study would compare these various scores against scores estimated by deadline methods (e.g., Lohman, 1989, 1991). A less expensive approach to validation is simply to determine which of these various scores has the best psychometric properties (internal consistency reliability, test-retest reliability, and factor validity).

Modeling

Assuming success in developing good proficiency scores for each of the tests, the second step is to evaluate alternative models of the variance-covariance matrix of those scores. Our initial model assumes that the common variance of each score can be partitioned into four sources: general (affecting all scores), process, domain, and paradigm factors. The error variance can be partitioned into occasion, measurement error, and unique (test-specific) sources. (Paradigm could alternatively be considered error, albeit correlated error since each paradigm is represented three times; it is a matter of semantics.)

There are actually many ways to implement this model as described. One is as follows. All scores load on the general factor (one factor). All scores except WM additionally load on a process factor (e.g., PS, FL, GK; five more factors total). We equate g with WM in this model to identify the model. Domain cannot be represented by three separate factors because that would give a linear dependency with the g factor. In analogy with dummy modeling in regression analysis, domain could be represented by two factors. S domain tests load on a spatial (as opposed to articulatory) factor (one more factor), and Q domain tests load on a quantitative knowledge (as opposed to verbal knowledge) factor (one more factor). This equates g (WM) with the verbal domain. There will also be 12 paradigm factors (number of paradigms in each cell minus 1, for identification purposes). There is no paradigm manipulation in the GK cells however. This gives a total of 19 common variance factors. All these factors are orthogonal to each other (the variance common to all scores is represented in the general factor), with the possible exception of the paradigm factors. They would be orthogonal to the g , process, and domain factors, but in some cases they could be correlated with one another. Each PS paradigm maps onto a corresponding WM paradigm, for example (this could be modeled hierarchi-

cally rather than as a correlation, however). The error variance factor modeling is straightforward and not worth elaborating here.

This degree of precision in variance partitioning on this heterogeneous a collection of test performances goes beyond anything ever attempted in psychology, to my knowledge. Consequently, virtually everything we learn in conjunction with fitting this model should be a contribution to the literature. Among the issues that can be examined are these: (a) how much of the variance is due to the general versus the process factors? (b) how much of the variance in working memory capacity is due to processing speed? (c) are the domain factors more important for some factors than for others, as the radex suggests (see earlier discussion)? (d) which factors (as opposed to tests) are most stable over occasions? (e) does it matter which paradigm one uses to measure the various factors? (f) which factor is more important to determining learning ability, knowledge or working memory capacity? (g) or is learning ability a very separate (independent) kind of ability? (h) is the procedural-declarative distinction justified?

Alternative Approaches

A criticism of Guilford's approach was that procrustean methodology, unlike confirmatory factor analysis (CFA), is weak in its ability to reject models and consequently exposes the analyst to hypothesis confirmation bias. I believe that confirmatory factor analysis methods are relatively immune to the flexibility criticism. Still, I concur with Glymour, Scheines, Spirtes, and Kelly (1987) that a more subtle confirmation bias can still exist when using CFA methods. Researchers typically evaluate two or three models derived from some "pet" theory or variants of that theory, and compare these to "null models" or to the theoretically perfect model (i.e., the one that fits the data to the limits of the stability or precision of the data). What CFA researchers (including myself) usually fail to consider is the legion of causal models that could have generated the data matrix as well as or even better than did the chosen model. Thus confirmation bias exists not in the statistical machinery, which is perfectly capable of rejecting inappropriate models, but in the choice by the researcher of which models to evaluate in the first place, and in the choice to avoid considering alternative models that might be unrelated to extant theory.

To loosen oneself from this kind of confirmation bias, it would be useful to conduct a somewhat serendipitous pursuit of alternative models for any correlation matrix, such as one generated by the CAM taxonomy data set. Specifically, it would be informative to compare a CAM model, such as the one discussed several paragraphs back, to models suggested by alternative approaches, such as exploratory factor analysis (Carroll, 1989) and TETRAD (Glymour et al., 1987). TETRAD is a computer program that builds multiple

causal models for a correlation matrix from scratch. It does this by searching the matrix for two kinds of correlation patterns, (a) zero semipartial correlations among three-variable triads, and (b) vanishing differences among four-variable tetrads. Both these patterns are consistent with certain causal models of the variables (for triads, A causes B causes C; for tetrads, a general latent factor affects all four variables).

It would also be useful to assess the generality of the CAM model by applying it to other data sets. For example, Wothke et al. (1990) have administered the ASVAB and the ETS kit of cognitive reference tests to a common subject pool. Exploratory factor analysis of their correlation matrix has yielded a 16-factor solution (Carroll, personal communication, November 1990). A question is whether a CFA based on CAM factors could provide an equally good fit to the data.

CAM TAXONOMY COMPONENTS AND SUBCOMPONENTS

Models based on the CAM taxonomy identify components at the level of taxonomy rows and columns, or possibly cells. But as I argued in an earlier section of this chapter, it may be fruitful to consider subcomponents, that is, multiple variance sources within each of the CAM taxonomy cells. In this section, I consider some of the major CAM row components (“cognitive processing factors”), review what we have learned about them, and then discuss possible subcomponent decompositions of those row components.

Working Memory (WM)

We have developed numerous, diverse working memory tests conforming to Baddeley’s (1986) definition that they “require the simultaneous processing and storage of information” and “measure various contents” (pp. 34–35). Although this definition is rather broad, it has turned out to be useful in developing tests, and in evaluating whether an existing test was a “good” measure (in the sense of high convergent and discriminant validity) of working memory capacity. The following is a summary of our main findings regarding working memory:

1. *WM tests, that is, those developed on the basis of Baddeley’s general definition, form a coherent factor.* That is, WM tests, despite being diverse in form and content, are more highly correlated with one another than they are with other kinds of tests, such as knowledge tests (e.g., vocabulary) or processing speed tests (e.g., arithmetic operations) (e.g., Kyllonen & Christal, 1990).
2. *WM tests are maximally valid,* that is, they correlate more highly with just about any kind of learning criterion than any other kind of test does. It

does not seem to matter much whether the criterion is initial learning, rate of improvement, or final learning level.

3. *WM capacity is an ability almost identical to general reasoning ability.* The only difference is that reasoning tests tend to draw a little more on (or, in a theoretically more neutral vein, be more highly correlated with) general knowledge (Kyllonen & Christal, 1990).

4. *WM tests systematically vary in validity.* The most valid WM indicator we have thus far developed is the *ABC numerical assignment* test (example item at the median difficulty level: $A = C + 3$; $C = B/3$; $B = 9$; $C = ?$ [3]; $A = ?$ [6]; $B = ?$ [9]). Other fairly good measures are *alphabet recoding* (example item: $J; P; C; + 2 = ?$ [L R E]); and *mental math*, when problems are fairly hard (e.g., $16 \times 12 = ?$ [192]); and *continuous paired associates* ($A = 14$; $B = 23$; $C = 6$; $D = 12$; $A = ?$; $A = 51$; $D = ?$; $D = 35$, etc.). All these tests show consistently high loadings on a working memory factor, regardless of what other tests are taken as indicators of the factor (Kyllonen & Christal, 1990; Pena & Tirre, 1990a).

5. *Simple span measures are valid*, certainly not as poor as Daneman and Carpenter (1980) suggest (see also Engle, 1988; Turner & Engle, 1989). Still, of the WM tests we have evaluated, the less valid ones are the *Span* measures, both simple span (digit, word) and sentence span, and *4-term ordering* (e.g., A precedes B; C follows D; Set 1 [A and B] does not precede Set 2 [C and D]; Which ordering is correct? [DCAB]). These tests tend to have higher uniquenesses (i.e., lower average correlations with other working memory tests), and higher correlations with the general knowledge factor than other working memory tests. The problem is that a non-working-memory component, specifically knowledge, affects people's scores. This can be demonstrated both correlationally and experimentally (Tirre & Pena, 1990).

What do these observations tell us? Perhaps most importantly, they argue against the idea that working memory capacity is strictly domain-specific (cf., Carpenter & Just, 1988). Certainly one can increase one's effective working memory capacity by increasing (strengthening and broadening) one's relevant knowledge base. But the coherency of a working memory factor, working memory's high correlations with performance on diverse learning tasks and reasoning tasks, and the finding that the best measures of working memory are those that are relatively independent of knowledge suggest this. It should be fruitful to investigate the components of a *general* working memory system; a system that is not domain-specific. Our findings also argue against the idea that working memory reflects multiple specialized capacities, as some have argued (Carlson, Khoo, Yaure, & Schneider, 1990). It may be useful to attempt to disentangle working memory subcomponents, but there is a strong suggestion in what we have found that working memory capacity may be the general (*g*) factor in cognition.

Thus far we have not systematically investigated possible subcomponents of working memory, but we have a variety of ideas on how to proceed in that direction. One proposal is to start with the components of Baddeley's model: The central executive, the visual-spatial scratch pad, and the verbal-articulatory loop might each have independent capacities. Also, the functions of the central executive—scheduler, strategy selector, information integrator, storage—might be independent sources of individual differences. In any event, a more refined WM model would enable the exploration of WM's relationship with other factors, such as reasoning ability, processing efficiency, and knowledge, in an attempt to pin down the locus of correlations we have noted in prior research. An independent research stream would involve an error analysis of working memory failures. It might be interesting to develop a taxonomy of error types and to identify task and individual characteristics that lead to particular errors. This work could establish bridges to cognitive task analysis.

Processing Speed (PS)

Much of the initial work on the LAMP project (1983–1985 period) centered on measuring processing speed. As discussed elsewhere (Kyllonen & Christal, 1989), there were three primary reasons: (a) computer testing made it feasible, (b) successful performance in many critical specialties (air crew, controller) obviously depended on processing speed, and (c) processing speed is a more sensitive variable than is accuracy—opening the possibility of identifying separable information-processing abilities more precisely than had been done in the past. Further, much of the cognitive and individual differences literature in the late 1970s and early 1980s focused on the measurement of processing speed. This made available a fairly sophisticated methodology for its measurement. And there was (and to a lesser extent there still is) considerable optimism that broad ability factors would eventually be understood in terms of more basic processes that were revealed through processing speed measurements. The following summarizes results from these investigations of processing speed:

1. *There is a general speed factor.* In countless published (Kyllonen, 1985) and unpublished studies, we have found that an exploratory factor analysis of accuracy and latency scores from a set of tests yields separate speed and accuracy factors at the highest order. We also have demonstrated that processing speed influences learning independently of knowledge level (Christal, 1990; Kyllonen et al., 1991), indicating that processing speed is more than simply a reflection of knowledge. And in other studies (Chaiken, 1990; Woltz, 1988) we have encountered the nuisance of method overlap: Latency measures reflecting one process correlate with latency measures reflecting supposedly independent processes even when care is taken to statistically remove obvious sources of

overlap, such as button-pressing latency. Considered together, these observations point to the existence of a general speed factor—a broad, independent source of individual-difference variance.

2. *On the other hand, the speed factor is not that critical to learning (at least the kinds we have looked at so far).* We have found that processing speed does not predict learning, once the effects of incoming knowledge and working memory capacity are statistically controlled for (Kyllonen & Stephens, 1990; Shute & Kyllonen, 1990). In the one study where we did find that processing speed predicted learning, we did not assess working memory (Kyllonen et al., 1991). Furthermore, processing speed predicted learning only when study time was extremely short. These observations suggest the following linkage: Processing speed determines working memory which determines learning; processing speed does not independently influence learning unless the learning task is one of high information flow.

3. *Latency scores are stable within a session; unstable over sessions.* We have found that internal consistency reliability of latency scores tends to be extremely high; generally above .95. Test–retest correlations tend to be lower; typically below .70. More complex processing tasks (e.g., semantic matching) are more stable than simple tasks (e.g., choice reaction time).

4. *Accuracy scores from processing speed tasks are more valid predictors of learning criteria than are speed scores.* We have found this result in a couple of studies. Christal (1990) identified a “processing accuracy” factor loaded by accuracy scores from processing speed tests (physical, name, and meaning identity, and choice reaction time), all of which had very high mean accuracy levels (ranging from 93% to 98% correct). This factor was found to be independent of both a working memory factor and a processing speed factor, and it correlated ($r = .20$ to $.30$) more highly with success on a procedural (logic gate) learning task than did the processing speed factor. Tirre (in press) found a similar result on a paired associates learning criterion. Tirre’s processing speed tasks were meaning identity, lexical decision, and an orthographic processing task (“Which sounds like a real word, lekchure or lakchure?”).

5. *Latency variance is more valid than latency means.* On choice reaction time tasks, one can compute, for each subject, mean RT and the variance of RT over trials. Jensen (1987) has shown that the variance measure often correlates more highly with an external criterion (such as a measure of general intelligence) than does mean RT. Fairbank, Tirre, and Anderson (1991) replicated Jensen’s finding on a wide variety of processing speed tasks.

6. *There is evidence for at least three fairly independent dimensions of processing speed.* Multidimensional scaling analyses of information processing tasks suggests two fairly independent dimensions of processing speed: the degree of perceptual processing involved in the task and the degree of memory search required (Kyllonen, 1985). Other research suggests an independent motor processing dimension (e.g., Ackerman, 1988). Coincidentally, this tri-

chotomy corresponds to the perceptual-semantic-motor breakout in cognitive architectures in the human performance literature (e.g., Card, Moran, & Newell, 1983).

7. *Latency is determined by two factors; carefulness and processing speed.* We have replicated a curious finding in a number of unpublished studies. On various (but not all) working memory tasks, we find a positive correlation between latency and accuracy: Those who take longer do better. Yet when we enter latency and accuracy in a regression equation predicting some third variable, such as a general ability composite, we get a positive regression weight for accuracy and a negative regression weight for latency. That is, at a given accuracy level, high-*g* subjects are faster. This pair of results suggests separable carefulness and processing speed factors. Carefulness (slow latency) determines accuracy on the task, but controlling for accuracy, processing speed (fast latency) is an independent predictor of external criteria.

Although these seven results are intriguing, our efforts admittedly encountered some obstacles—the first, methodological, the second more conceptual. The methodological problem was speed–accuracy tradeoff. Comparing response time for individuals who are responding at different accuracy rates is problematic to say the least. The conceptual problem concerned disentangling a speed factor per se from other influences on response time, such as knowledge. One can respond quickly because the knowledge required to respond is well practiced. But to what factor should speedy responses then be attributed, knowledge or processing speed? And is this distinction psychologically fundamental or is processing speed simply a convenient dependent variable, an indicator of something, such as knowledge, but not a source of individual differences in its own right?

One idea is that for any task, there are two determinants of latency: processing speed and persistence (or carefulness). A goal would be to develop a methodology for separating them. I suspect that many of the confusions and inconsistencies in our previous investigations of processing speed (e.g., observations 3, 4, and perhaps 2) have resulted from our confounding the two components. For example, it could be that processing speed on any task is a fairly stable factor, but that carefulness fluctuates by the hour. I see the settling of the speed–accuracy tradeoff problem as a prerequisite to further work on processing speed, such as the identification of subcomponents.

But assuming we can resolve the speed–accuracy problem, a question is whether there is a single processing speed factor or whether it is more appropriate to posit multiple dimensions of processing speed. The generic information processing model and previous empirical investigations (Kyllonen, 1985) converge in suggesting three processing cycles: one each corresponding to perceptual encoding, semantic retrieval, and response execution. There may be justification for a further decomposition. Perceptual encoding can be broken down into attention orienting and attention shifting; semantic retrieval can be

broken down into transformation, matching, and decision phases; and response execution can be broken down into initiation, movement, and completion phases. A second and orthogonal components-of-processing speed distinction is designed to account for a ubiquitous characteristic of response time distributions not yet discussed—their inevitable positive skewness. Deadline methodology yields a comparable result: Accuracy (percent correct) over systematically increasing deadlines shows negatively accelerated growth as a function of exposure. Following Pieters and van der Ven (1982), one might suspect that the processing cycle (and it could conceivably be any of the processing cycles) could be divided into on versus off time. If so, this might explain observation 5, at least partially.

Declarative Learning (DL)

We initially treated the ability to learn as an ability to be accounted for by other abilities (e.g., working memory capacity, processing speed) rather than as a unique ability in its own right. However, in several studies on procedural learning (Kyllonen & Stephens, 1990; Shute & Kyllonen, 1990), we found that learning the propositional foundations of some skill (i.e., declarative learning) predicted subsequent acquisition of the skill itself (i.e., procedural learning), controlling for other cognitive factors. An obvious follow-up question is whether the fact that associative learning predicts procedural learning is due to item effects (strong items are more easily proceduralized) or a mechanism effect (associative learning ability per se is related to procedural learning ability per se). To test this question it is necessary to measure associative learning on one set of items, and procedural learning on another set. Consequently, we have developed “stand-alone” associative learning tests that we now can administer in procedural learning studies. The following is a summary of some of our key findings regarding associative learning ability.

1. *There is a general associative (declarative) learning factor.* The psychometric literature suggests a general associative memory factor (the *Ma* factor; French, 1951; Thurstone, 1938), which we and others have verified (Kyllonen & Tirre, 1988; Tirre, 1990; Underwood, Boruch, & Malmi, 1975). The key finding is that the factor is quite broad (i.e., general). Most of the variance in performance on quite diverse associative learning tasks (study test, trials to criterion, paired associates, free recall), is determined by a general associative learning ability rather than by any paradigm-specific factors.

2. *Associative learning ability is moderately related to working memory capacity and to general knowledge, and somewhat less related to processing speed.* We have found correlations in the .30s to .50s range between associative learning ability, working memory capacity and general knowledge factors (Kyllonen et al., 1991; Tirre, in press). Note that these are far from perfect correlations, suggesting that there is an associative learning ability independent of

other cognitive factors. Associative learning ability is related to processing speed only when study time is short. We have also found that knowledge of mnemonics, while important in determining performance on certain associative learning tasks, is independent of associative learning ability per se (Kyllonen et al., 1991), and thus is essentially a nuisance factor.

3. *Decay over time is not distinguishable from general associative learning ability.* It has been claimed that there are no individual differences in retention (Shuell & Keppel, 1970; Underwood, 1954). However, we have found individual differences in retention (Kyllonen & Tirre, 1988; Woltz, 1990a; Woltz & Shute, 1991), a discrepancy due to how individuals are equated for initial learning. However, even these individual differences appear to be predictable from initial learning rate, and thus there is yet little evidence for a decay factor independent from a learning factor.

4. *Resistance to interference (susceptibility to fan) is not a very reliable individual differences dimension.* There also is scarce evidence for a resistance to interference (RTI) factor independent of general learning ability. Chaiken (1989) estimated an RTI factor on a "fan task," but found that such a factor had fairly low reliability ($r_{-xx'} = .30$).

5. *Incidental learning strongly predicts intentional learning.* Intention does not affect learning (Anderson, 1985). An individual differences corollary would be that intentional and incidental learning are the same ability, a prediction supported by some research (Tirre, 1990). Tirre had subjects verify (true/false) sentences such as "plumbers work with pipes," then, after 50 such trials, gave them a cued recall test (e.g., plumbers = ? [pipes]). He created two versions of the task, the only difference being that in the version administered first (the incidental learning version), subjects were not told they would be tested, whereas in the version administered next (intentional learning version), subjects were told. The correlation between the two versions ($r = .70$) was close to the reliability of the two tests.

The existence (observation 1) and importance (observation 2) of a general associative learning factor seems clear. But our (and others') past research is limited in a number of ways. First, we have focused most attention on the initial stage of associative learning, the probabilistic transfer of a trace to long-term memory on initial exposure. Models of associative learning from the cognitive literature (e.g., Anderson, 1976) posit a number of additional independent parameters governing declarative learning besides *transfer to long-term memory* (p), most notably *strengthening* (s), *decay* (d), and *interference*, or *fan* (f). The individual-differences question is whether these parameters correspond to independent abilities. That is, do those who initially learn a fact easily (i.e., high p subjects) also benefit most from repeated exposures (high s subjects), experience the least decay over time (low d subjects), and suffer least from interference from related material (low f subjects)? Or are these learning parameters uncorrelated?

A second limitation to our past research is that we have focused on a rather narrow range of associative learning tasks, particularly, paired associates tasks with verbal materials. Examining other kinds of learning tasks and other kinds of stimuli, such as figural stimuli, might enable identifying parameter abilities (viz., decay [d], and resistance to interference, or fan [f]) that we have failed to clearly identify in our previous work with verbal stimuli (observations 3 and 4). We also suspect that besides the nature of the stimuli, another reason we may have failed to identify a clear decay factor is that we have focused on short-term retention (less than 3 hours), primarily for pragmatic reasons.

Procedural Learning (PL)

We have conducted a number of studies investigating correlates of the ability to acquire cognitive skill in procedural learning tasks. These include logic, computer programming, electronics troubleshooting, and if-then decision making. In these studies, we treated procedural learning as the dependent variable, the variable to be accounted for by other cognitive factors (Chaiken, 1990; Kyllonen & Stephens, 1990; Shute & Kyllonen, 1990; Woltz, 1988). However, as with declarative learning, an obvious prior question is whether there exists a general procedural learning ability at all, that is, a common thread running across diverse cognitive skill-acquisition tasks. If so, that would suggest that general procedural learning ability ought to be treated as an independent variable (perhaps measured by a battery of short, procedural learning tasks) in studies of the determinants of learning on specific procedural tasks. Treating procedural learning as an independent variable (or set of variables) would enable conclusions such as "what determines how quickly one will learn to program a computer is (a) the degree to which one possesses general procedural learning ability and (b) specific knowledge about mathematics concepts," for example. The question of the determinants (or subcomponents) of general procedural learning would still be an interesting, but independent question.

On the other hand, if there were no general procedural learning factor, then, of course, procedural learning ability could not be a predictor of success on particular procedural learning tasks, such as programming. There are two ways in which a procedural learning factor could fail to materialize. One is that procedural learning could be found to be highly task specific, a conclusion hinted at by some research on psychomotor learning (Ackerman, 1987, 1988). Or, procedural learning could be found to be indistinguishable from general associative learning (or, obviously, a little bit of both). We have not conducted definitive research yet on these issues, but what follows summarizes what we have found so far.

1. *There may or may not be a general procedural learning factor independent of declarative learning ability.* In a series of studies (Kyllonen & Stephens, 1990; Shute, 1991; Shute & Kyllonen, 1990; Soule, 1990), we taught

various cognitive skills (logic gates, programming, graph reading) by first teaching the declarative facts underlying the skill, then providing practice solving problems in the skill domain. Our initial expectation was that performance on the declarative part would be predictive of performance on the later procedural (problem-solving) part, but we have been quite surprised at the degree to which declarative acquisition predicted procedural acquisition. Depending on the study, and psychometric details of how procedural learning is operationalized, we have found that either there is no procedural learning factor independent of declarative learning, or that such a factor is fairly weak, that is, not uniquely accounting for much of the variance on the procedural learning portion of the task.

2. *Early procedural learning is strongly dependent on working memory capacity.* Perhaps one of the reasons for our failure to identify a general procedural learning factor is that procedural learning, at least initially, is entirely a function of working memory capacity. We (Kyllonen & Stephens, 1990; Shute, in press; Shute & Kyllonen, 1990; Woltz, 1988) and others (Ackerman 1988) have found that working memory capacity is highly correlated with initial procedural learning.

3. *Late procedural learning may be governed by factors other than working memory capacity, such as processing speed and a memory strengthening ability.* It may be that procedural learning only becomes a unique ability after individuals have gotten past the initial, working-memory intensive phase of learning. For example, some research has shown that processing speed becomes the factor governing performance after subjects are reasonably well acquainted with the task (Ackerman, 1986, 1988; Chaiken, personal communication, December 1990). Other research has shown that a kind of memory strengthening ability governs later learning (Woltz, 1988).

It seems that the best way to clear the emerging confusion regarding a procedural learning factor is to focus first of all on its theoretical similarities to and differences from a declarative learning factor. In particular, what declarative and procedural learning might have in common are the *p*, *s*, and *d* parameters. What seems to be unique to procedural learning is the importance of grouping related problem-solving activities into unified sequences in memory—a process that has variously been called chunking, composition, or the creation of “macro-operators.” It would be useful to seek to identify a chunking factor in procedural learning, then to determine the correspondence between declarative and procedural learning parameters.

It may be that we will not make much progress in identifying a procedural learning factor until we have a solid theoretical foundation for describing procedural learning. To date, our thinking on procedural learning, and the way in which we have sought a procedural learning factor, has been heavily influenced by Anderson’s (1983, 1987) ACT* theory. But recently, a competing “instance theory” of procedural learning has been articulated (Logan, 1988).

Although both theories can account for the major findings in procedural learning, the theories are expressed in quite different ways. If Logan's theory turns out to be a better description of procedural learning phenomena, we might have to rethink our approach to identifying procedural learning factors.

Declarative Knowledge (DK)

We have conducted numerous investigations of the influence of knowledge in learning and performance. Following is a summary of our main findings.

1. *There is a strong easy-to-measure general declarative knowledge factor.* We have found that scores on standard vocabulary tests and general knowledge surveys are highly intercorrelated, often to the limits of their reliability. Further, informal (unpublished) item analyses of such tests generally yield either a single factor solution or a multifactor solution in which all but the first factor are weak and mostly uninterpretable.

2. *General knowledge predicts learning (independent of specific knowledge).* In almost every cognitive correlates study we have conducted, we find that general knowledge predicts learning, controlling for other cognitive factors, such as working memory capacity, processing speed, and general learning ability. Presumably, the knowledge base provides the essential material to which new knowledge is "tied." A dense compared to an impoverished base provides more material to which new material can be related (Kyllonen et al., 1991). The perhaps unexpected finding is that the possession of broad general knowledge predicts learning after controlling for the amount of domain-relevant, specific knowledge (Shute & Kyllonen, 1990).

3. *Specific knowledge predicts learning (independent of general knowledge).* In a series of studies of significant learning (learning tasks that take several days), we have found that knowledge of domain-relevant concepts (e.g., in computer programming, knowledge of the concepts "string," "integer," "variable," etc.) affects both initial learning and how much is ultimately learned.

4. *Specific knowledge predicts effective working memory capacity and effective processing speed in the specific domain.* It is fairly simple to show that if subjects are exposed to a particular fact, they will perform better on either a processing speed test (Woltz, 1990b) or a working memory capacity test (Pena & Tirre, 1990) that involves that fact. The expertise literature also seems to suggest that cognitive feats involving tremendous memory capacity and lightning-fast access to knowledge (i.e., improved processing speed) can be traced to extensive domain-specific knowledge. An interesting aside is that these feats do not necessarily result from high general knowledge, nor do they transfer to domains outside the specific area of training.

There clearly is a general knowledge factor, and that factor is an important determinant of learning (observations 1 and 2). The looming question is

whether there are additional specific knowledge factors, such as science knowledge, social knowledge, or knowledge of history, independent of general knowledge. On the one hand, we have observed specific knowledge effects (observations 3 and 4). On the other hand, we have failed to identify specific factors from analyses of general knowledge surveys (observation 1). We can design specific knowledge tests, but they suffer from an ad hoc quality. What would be most useful would be a broad knowledge taxonomy, which could serve as the basis for the development of principled specific knowledge tests.

The current CAM framework contains a very simplistic knowledge taxonomy—verbal, quantitative, and spatial. It clearly would be worthwhile to elaborate on this tripartite taxonomy, perhaps beginning with the verbal portion of it. For example, if one could generate numerous, diverse knowledge items, it would be interesting to note whether items cluster into interpretable specific knowledge groups, such as science, math, and so on. The obvious (but incorrect) way to test this would be to compute a correlation matrix over subjects from question response data and conduct a factor or cluster analysis on that matrix. Unfortunately, correlations do not adequately reflect question similarity, for various technical reasons. In particular, questions can be similar (or dissimilar) because either they are at the same (different) difficulty level or because they tap the same (different) specific knowledge, factors we would have to be able to separate before we could identify specific knowledge factors. Thus the question similarity matrix will have to be something other than a correlation matrix. Similarity indices other than correlations have been developed in the occupational survey literature. There they must determine the similarity between any two jobs on the basis of overlap in tasks performed on the job, whereas we are trying to determine the similarity between questions on the basis of overlap in who successfully answers those questions.

Procedural Knowledge (PK)

Procedural knowledge (know how) is knowledge of how to solve problems. It ranges from very specific (how to solve this particular problem) to very general (how to go about solving never-before-encountered problems). The imparting of specific problem-solving skill is the goal of training, and thus we have focused on the assessment of very general procedural knowledge, which should be more amenable to testing. Besides, according to current conceptions (e.g., Anderson, 1983; Langley, 1985; Newell, 1990) the path to specific procedural knowledge begins with general procedural knowledge, which is transformed with experience. Thus assessing general procedural knowledge seems like the right place to focus attention. Unfortunately, there is a problem with measuring procedural knowledge—by definition it is not readily accessible (Squire, 1986). Thus, unlike with declarative knowledge, where one can simply ask a person whether they know this or that fact, with procedural knowledge, this approach

will not work. We cannot assume that people can easily talk or introspect about their problem-solving knowledge. We can only measure procedural knowledge indirectly, by having people solve problems.

In our past research (e.g., Kyllonen & Stephens, 1990), we have found it quite difficult to identify a procedural knowledge factor. Either the factor is indistinguishable from working memory capacity, or, if it can be distinguished, it is not uniquely related to procedural learning. At this point, we can either conclude that there is no such an entity as a general procedural knowledge factor, or that we have simply failed to test for it properly. Given the prominent role general procedural knowledge plays in theories of cognition (e.g., Anderson, 1983; Newell, 1990), we are willing to accept that we simply have not tested for it properly.

To establish whether there is a general procedural knowledge, I think a fruitful path would be to administer batteries of novel problem-solving tasks arising from three disparate bodies of literature, and look for unities in performance. The first group of tasks, developed from the psychometric literature, are the inductive reasoning (I) tests. These are the standard series, sets, and matrices tests. The second group of tasks comes from the information processing literature which has sought an explanation for how people (or how machines can be made to) solve novel problems for which algorithms do not yet exist. The belief is that people rely on "weak methods," very general problem-solving strategies that can be applied to a wide variety of situations, but that are slow, effortful, and error prone. The tasks typically employed in these kinds of investigations include the missionaries-and-cannibals task and the tower-of-Hanoi problem. The third group of tasks are insight tasks (Davidson & Sternberg, 1984; Sternberg, 1986), of the kind one typically finds in problem-solving puzzle and game books. These differ from other problem-solving tasks in being accompanied by what Metcalfe (1986) has called "a subjectively catastrophic insight process," in which subjects do not experience a gradual closing in on a solution, but rather a dramatic leap between being in a subjectively "not even close" state to a solution state in a short period of time.

NEW COMPONENTS

The CAM taxonomy includes rows for tasks on which we have not yet systematically collected data. These represent tentative additions to the CAM taxonomy. In this section I discuss the empirical and theoretical motivations for these additions.

Temporal Processing (TP)

Recently, there have been a number of investigations conducted both in our laboratory (Yee, Hunt, & Pellegrino, 1991; Snow, Chastain, & Jackson, 1992)

and elsewhere (Irvine, Wright, Dennis, & Gould, 1991) that have examined the ability to track objects dynamically in space. Two kinds of tasks have received most of the attention, which I will refer to as the "horse race" and "intercept" tasks. In the horse race task, two objects are shown moving across the screen in the same direction along parallel paths. They disappear behind a wall, and the subject is to estimate which object will be first to reach a finish line. To make the task challenging, the two objects can move at differing rates, but the slower one will start closer to the finish line to make it a close race.

In the intercept task, two objects are shown moving toward each other on intersecting paths (perpendicular or oblique), and the subject is asked to determine whether they will collide or miss each other. Or, in a related version, the subject will have control over the release of the second object, and the task is to release the object at the proper time so as to intercept the first object. In both these kinds of tasks, various features can be manipulated (e.g., speed of the objects, relative speed, distance to target, length of the occluding wall), and there has been some work on determining which are most critical in determining task difficulty.

There is some preliminary evidence that the ability to perform well on the horse race kind of task (and its variants) is correlated with the ability to perform well on the intercept kind of task (and its variants), indicating that there might be a general "dynamic spatial ability." Interestingly, this dynamic spatial ability is somewhat independent of the ability to perform well on standard (static) tests of spatial ability (Hunt, Pellegrino, Frick, & Alderton, 1988). From casual analyses, it seems that a number of critical specialties (e.g., pilot, controller) routinely require the exercise of a dynamic spatial ability, and thus it seems important to investigate this ability in more depth.

From informal observation, I believe there might be a correspondence between dynamic spatial tasks and a class of tasks, which I will refer to as counting tasks, that require brief-duration time keeping. An example of a counting task is this: Subjects are shown a sequence of ascending digits (1, 2, 3, ...), each of which is presented for only a brief period of time (e.g., 200 msec, with, say a 200 msec blank between each digit). The effect is similar to observing the tenths counter on a digital stop watch, although not as fast. After the sequence reaches 40 the screen goes blank. The subject's task is to press the space bar when the sequence would be expected to reach 100. The task seems similar to the dynamic spatial tasks—both involve estimating when the object (the spatial object or the counter) will hit a "finish line," and in both, the object is occluded during the final half or so of its "run." The difference is that there is no spatial processing involved in the counter task.

It is interesting to imagine counting tasks that are isomorphic to the dynamic spatial tasks. For example, a counting isomorph to the horse race task would present two counters, one perhaps starting at 20, the other at 0, and they would count at different rates (the 0 counter moving faster). Both counters would

blank at, say, 50, and the task would be to indicate which counter would reach 100 sooner. A counting isomorph to the intercept task might involve starting a slow counter (from 0), then allowing a subject to start a faster counter (also from 0) at such a time as to force a “collision” (i.e., a point at which both counters coincided at the same number in some range (say, 60 to 80). Because it is so easy to imagine counting-task isomorphs to dynamic-spatial tasks, I suspect there very well might be a general “temporal processing” ability.

Time-Sharing Ability (TS)

The notion of a general time-sharing ability has considerable intuitive appeal. The idea is that some individuals, those with ample time-sharing ability, might be able to perform two tasks simultaneously, and suffer little deterioration in their performance on either task. Others, while equally capable of performing the two tasks when performing them separately, might “fall apart” when having to do them at the same time. More formally, the time-sharing ability hypothesis is that performance on some kind of combined task is independent of (unpredictable from) performance on the two tasks making up the combined task when each of them were performed separately (controlling for task reliability, and ideally, task uniqueness). As was the case with dynamic spatial ability, job-task analyses suggest that certain critical specialties (e.g., pilot and controller) might heavily tap such an ability, if it exists.

Unfortunately, there is little, if any evidence for the existence of a time-sharing ability. Some of the older literature has been justly criticized on methodological grounds (Ackerman, Schneider, & Wickens, 1984). One problem is that researchers have failed to examine whether there is any unique variance in the combined task performance, that is variance unrelated to single task variance. Even when this has been done, a second problem is that researchers have uncritically accepted uniqueness on the combined task (that portion of combined task variance that is unpredictable from constituent single-task performances) as evidence for a time-sharing factor. In fact, there are many potential reasons for task uniqueness (e.g., unreliability, subject-to-subject variability in allocation strategy). To demonstrate the existence of a *factor*, it is necessary to show that the variance unique with respect to the constituent single task performances is common with something else—ideally a “uniqueness” from a second combined task performance.

A distinction is sometimes made in the literature between dual tasks and coordination tasks (Pellegrino, Hunt, & Yee, 1989). A dual task is one in which an individual performs two independent tasks; in some cases one of these is designated a primary task, the other a secondary task, but such a designation is not critical. A major problem with dual tasks is scoring them. Subjects might differ in their attention allocation strategy, so that some devote more attention to task A, others to task B. At that point, the scoring system

will determine who performed better on the combined task: If the scoring system weights task A performance more heavily, those who devoted more attention to task A will benefit and will appear to have performed better on the combined task. But this is clearly a scoring artifact as opposed to an unbiased indication of better performance. In this regard, it is interesting that at least one good study of time-sharing, one that avoided the obvious methodological problems listed in the previous paragraph, identified what was labeled a time-sharing factor. The problem is that the factor was identified by only one (of four investigated) constituent tasks, and was identified by both positive and negative loadings, indicating that the factor merely reflected subjects' proclivity to allocate their attention either toward or away from the particular constituent task (Brookings, 1990).⁴

An alternative to the dual task is what Pellegrino et al. (1989) call the coordination task. A coordination task is one in which the constituent component tasks are logically related (rather than logically independent), so that both tasks must be performed accurately for the combined task to be correctly performed. An example is a horse race task (see "Temporal Processing" discussion) in which the subject indicates which object will reach the finish line first by providing a true or false verification to a sentence (task 2) such as "the black object will be beaten by the white object." Pellegrino et al. (1989) recommend the investigation of coordination tasks over dual tasks on ecological validity grounds, but perhaps an even greater benefit is that the problem of differential attention allocation is greatly reduced with coordination tasks. With a coordination task, the issue of whether to allocate more attention to task A or B is not simply a capricious choice for a subject, but rather is integral, in a way transparent to the subject, to overall success.

The ambiguity of past research on time sharing may well be the result of methodological shortcomings, including the use of dual rather than coordination tasks, and the statistical problems. But these now can be overcome, and I believe that time-sharing should be investigated as an ability. Selecting cognitive tasks from a comprehensive taxonomy, such as the CAM taxonomy, would be a good place to begin.⁵ It still might turn out that there is no general time-sharing factor, but if so, its dismissal could be on the basis of theoretical rather than methodological or statistical grounds. In particular, a reason one might not expect there to be a time-sharing factor is that time sharing (or, more generally, resource allocation) is putatively one of the functions of the central executive of working memory, according to Baddeley's (1986) model. Therefore, according to the model, time-sharing ability is a determinant of

⁴This is not actually a criticism of the Brookings study; Brookings seemed to have been well aware of this problem with his "time-sharing" factor.

⁵Studies conducted have not been atheoretical. For example, many have made use of Wickens' (1980, 1984) attentional resources taxonomy.

working memory capacity, which would make it difficult to identify independently of the working memory factor.

Interaction Factors (IN)

One kind of interaction factor is a time-sharing factor which I just discussed. Another kind of interaction factor is a product of two other factors. This is analogous to a product variable, such as $X_1 * X_2$ included in a regression model, such as $Y = b_1X_1 + b_2X_2 + b_3(X_1 * X_2)$. To identify an interaction factor, one would augment the variable matrix with multiple product variables from the indicators of two (or more) factors. For example, if $X_1, X_2,$ and X_3 were indicators of Factor 1, and X_4, X_5 and X_6 were indicators of Factor 2, then $(X_1 * X_4), (X_1 * X_5), \dots, (X_3 * X_6)$ would be the nine indicators of the product factor, say Factor 3. It is empirical whether the variance of Factor 3 is reliably greater than zero. If it is greater than zero, a product factor has been identified. Going back to the regression model, this would be analogous to finding that the regression weight for the interaction variable is significantly different from zero.

One can imagine natural product factors. For example, there might be a modest correlation between a processing speed and a working memory factor, but that correlation could mask some nonlinearities in the relationship. Consider this: To a point, greater capacity could allow for faster processing, but after that point, extra capacity could result in no additional benefits in processing speed (or even slower processing) because the working memory system would have to perform more maintenance activities. This kind of relationship could be captured by a product factor. We have not investigated factor interactions such as these, but they would be fairly simple to investigate because they involve no additional data collection, only additional analyses. An intriguing empirical issue is whether factor interactions, if they exist, are independent of corresponding time-sharing factors. Indeed, in any evaluation of the existence of a time-sharing factor, one should rule out the hypothesis of a corresponding factor interaction. To my knowledge, this has never been done.

Other Factors

In the cognitive literature, one finds many interesting tasks or theoretical constructs that have not been examined by individual-differences researchers. A question is whether these tasks or constructs qualify as components in the sense used in this chapter—as sources of individual differences variance. In this section I review some of these tasks and constructs that I believe may prove to be components, or at the very least, are worth examining in that respect.

Probabilistic Category Learning. In probabilistic category learning tasks, subjects must learn to associate exemplars (or exemplar configurations) to cat-

egories (e.g., symptoms to diseases) where the relationship of exemplars to categories is only probabilistic (“fuzzy”) rather than exact. An intriguing feature of this kind of task is that it bears at least a superficial similarity to lots of everyday learning and performance tasks. For example, troubleshooters learn to associate states of the machine to particular fault diagnoses, and what makes troubleshooting hard is that such associations are often probabilistic rather than exact. There are now models in the cognitive literature that suggest that probabilistic category learning is essentially associative (Gluck & Bower, 1988); the individual-differences prediction that we will test is whether the seemingly inductive ability involved in category learning is really nothing more than simple associative ability, or whether it reflects a unique mental skill.

Implicit Learning. Implicit learning tasks are ones in which learning is measured not by one’s ability to consciously recollect what was studied, but by one’s improved performance on some test that uses what was studied in some indirect way (e.g., word fragment completion). The literature is replete with demonstrations of the disassociation between regular explicit learning measures and implicit learning measures (Schacter, 1987), but it is probably still an open question whether implicit learning tasks are simply more sensitive memory measures rather than qualitatively different ones. The fact that performance on implicit learning tasks seems to be uniquely predictive of later procedural learning (Chaiken, 1990; Woltz, 1988) suggests that implicit learning might be qualitatively different from explicit learning.

Decision making under speed stress. From personal experience I have noticed that when tasks have fairly strict time deadlines they seem to induce mild emotional stress. We occasionally administer tasks with deadlines to minimize variance in response time (presumably, everyone will wait until the deadline before responding), and maximize variance in accuracy, as a way of bypassing interpretation difficulties stemming from the use of two dependent variables (i.e., latency and accuracy). It might turn out that this manipulation actually taps an independent ability (decision making under speed stress) rather than simply serving to put all the performance variance on the accuracy variable. We have not examined this effect, but it might be fairly easy to with data already in hand.

SUMMARY

The CAM framework is definitely work-in-progress, rather than a fully articulated “theory” of individual differences in cognition. Nevertheless, I and my LAMP colleagues intend to rely heavily on the CAM framework in conducting basic research on the identification and assessment of cognitive components over the next several years. It is useful to divide this research into two parallel activities which we hope will simultaneously inform and benefit from each other.

One activity, centered around the CAM Taxonomy as a whole, is concerned with evaluating global models of the relationships among ability components through confirmatory factor analysis. Alternative models can be constructed on the basis of the CAM taxonomy framework, and compared using data from the particular tests that have been developed to fit into CAM taxonomy cells. At any given time, the taxonomy, along with its associated tests, accompanied by models of their interrelationships, stands as our current view on how abilities are organized and on how they can be measured. The taxonomy serves as both an outline for a theory of abilities and as a sampling plan from which test batteries can be constructed. As new data and new analyses come in, we expect significant modifications to the CAM taxonomy, its models, and its measures.

The other activity is concerned with refining the ways in which we measure CAM taxonomy abilities. Here the goal is to develop and evaluate specific psychological tests. We do this by conducting correlational analyses of those tests or by testing detailed models of the information-processing activities involved in test performance. A result of studies conducted under this activity will be the nomination of particular tests administered in particular ways for inclusion in future versions of the CAM battery. This is why I stated in the very beginning of this chapter that research directed toward the construction of a test battery does not have to be atheoretical.

It is reasonable to expect both basic and applied benefits resulting from a program of research based on a theoretical framework such as CAM. The basic research benefit is that the framework is a necessary first step toward a model or theory for characterizing human intelligence, that is, individual differences in cognition. Along the way, we might also expect to improve our understanding of particular CAM components, such as working memory capacity, procedural learning, and the organization of declarative knowledge. Applied research benefits include the specification of a framework for aptitude testing (e.g., the CAM taxonomy), along with a battery of validated aptitude tests (the CAM battery). We might also anticipate additional spinoffs in the form of potential applications in the areas of performance assessment, training-embedded testing, and cognitive task analysis.

REFERENCES

- Ackerman, P. L. (1986). Individual differences in information processing: An investigation of intellectual abilities and task performance during practice. *Intelligence, 10*, 101-139.
- Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin, 102*, 3-27.
- 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., Schneider, W., & Wickens, C. D. (1984). Deciding the existence of a time-sharing ability: A combined methodological and theoretical approach. *Human Factors*, 26, 71–82.
- Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Anderson, J. R. (1985). *Cognitive psychology and its implications (2nd ed.)*. New York: Freeman.
- Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review*, 94, 192–210.
- Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Clarendon Press.
- Bordelon, V. P., & Kantor, J. E. (1986). *Utilization of psychomotor screening for USAF pilot candidates: Independent and integrated selection methodologies (Tech. Rep. AFHRLTR-86-4)*. Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Brookings, J. (1990). A confirmatory factor analytic study of time-sharing performance and cognitive abilities. *Intelligence*, 14, 43–59.
- Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Erlbaum.
- Carlson, R. A., Khoo, B. H., Yaure, R. G., & Schneider, W. (1990). Acquisition of problem-solving skill: Levels of organization and use of working memory. *Journal of Experimental Psychology: General*, 119, 193–214.
- Carpenter, P. A., & Just, M. A. (1988). The role of working memory in language comprehension. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing; The impact of Herbert Simon*. Hillsdale, NJ: Erlbaum.
- Carretta, T. R. (1989). USAF Pilot selection and classification systems. *Aviation, Space, and Environmental Medicine*, 60, 46–49.
- Carroll, J. B. (1989). Factor analysis since Spearman: Where do we stand? What do we know? In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology: The Minnesota symposium on learning and individual differences*. Hillsdale, NJ: Erlbaum.
- Carroll, J. B. (1990, April). Factors in the kit of cognitive reference tests. In R. Ekstrom & I. Bejar (chairs), *New directions for the kit of cognitive reference tests*. Symposium conducted at American Educational Research Association Annual Convention, Boston, MA.
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston, MA: Houghton Mifflin.
- Chaiken, S. (1989, April). *Resistance to interference: A measure of learning efficiency or skill fluency*. Paper presented at the American Educational Research Association Annual Convention, San Francisco, CA.
- Chaiken, S. (1990). *Test proximity effects in a single-session study of learning and individual differences: The case of activation capacity*. Brooks Air Force Base, TX: Air Force Human Resources Laboratory.
- Christal, R. E. (1989). *Estimating the contribution of experimental tests to the Armed Services Vocational Aptitude Battery (Tech. Paper No. AFHRL-TP-89-30)*. Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

- Christal, R. E. (1990). *Comparative validities of ASVAB and LAMP tests for logic-gate learning* (unpublished manuscript). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Cohen, P., Cohen, J., Teresi, J., March, M., & Valez, C. (1990). Problems in measurement of latent variables in structural equations causal models. *Applied Psychological Measurement, 14*, 183–196.
- Cronbach, L. J. (1990). *Essentials of psychological testing* (3rd/5th ed.). New York: Harper & Row. (Original work published 1979).
- Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466.
- Davidson, J. E., & Sternberg, R. J. (1984). The role of insight in intellectual giftedness. *Gifted Child Quarterly, 28*, 58–64.
- Dixit, S. (1991). *Factor structure of learning styles questionnaires* (unpublished manuscript). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Engle, R. (1988). *Working memory: An individual differences approach* (AFOSR Tech. Rep.). Charleston: University of South Carolina.
- Fairbank, B. A. (1988, February). *The application of nonstandard analysis techniques to cognitive data*, (unpublished manuscript). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory. Unpublished manuscript.
- Fairbank, B. A., Jr., Tirre, W. C., & Anderson, N. S. (1991). Measures of thirty cognitive tasks: Analysis of reliabilities, intercorrelations and correlations with aptitude battery scores. In P. L. Dann, S. H. Irvine, & J. M. Collis (Eds.), *Advances in computer-based human assessment* (pp. 51–102).
- French, J. W. (1951). The description of aptitude and achievement tests in terms of rotated factors. *Psychometric Monographs (No. 5)*. Chicago, IL: University of Chicago Press.
- Gagne, R. M. (1965). *The conditions of learning*. New York: Holt, Rinehart, & Winston.
- Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning Memory and Cognition, 16*, 430–445.
- Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology; General, 117*, 227–247.
- Glymour, C., Scheines, R., Spirtes, P., & Kelly, K. (1987). *Discovering causal structure; Artificial intelligence, Philosophy of science, and statistical modeling*. San Diego: Academic Press.
- Guilford, J. P. (1967). *Nature of human intelligence*. New York: McGraw-Hill.
- Guilford, J. P. (1982). Cognitive psychology's ambiguities: Some suggested remedies. *Psychological Review, 89*, 48–59.
- Guilford, J. P. (1985). The structure of intellect model. In B. B. Wolman (Ed.), *Handbook of intelligence, Theories, measurements and applications* (pp. 225–266). New York: Wiley.

- Gustafsson, J. E. (1989). Hierarchical models of abilities. In R. Kanfer, P. Ackerman, & R. Cudeck (Eds.), *Learning and individual differences: Abilities, motivation, and methodology*. Hillsdale, NJ: Erlbaum.
- Horn, J. L., & Knapp, J. R. (1973). On the subjective character of the empirical base of Guilford's structure of the intellect model. *Psychological Bulletin*, *80*, 33-43.
- Humphreys, L. G. (1976). A factor model for research on intelligence and problem solving. In L. B. Resnick (Ed.), *The nature of intelligence*. Hillsdale, NJ: Erlbaum.
- Hunt, E., Pellegrino, J. W., Frick, R. W., & Alderton, D. L. (1988). The ability to reason about movement in the visual field. *Intelligence*, *12*, 77-100.
- Irvine, S. H., Wright, D. A., Dennis, I., & Gould, P. (1991). *Time estimation, aiming, and tracking: Measurement of individual differences in the performance of military tasks* (Human Assessment Laboratory Tech. Rep. No. HAL3-91 [APRE]). Plymouth, UK: University of Plymouth.
- Jensen, A. R. (1987). Process differences and individual differences in some cognitive tasks. *Intelligence*, *11*, 107-136.
- Johnson-Laird, P. N. (1985). Deductive reasoning. In R. J. Sternberg (Ed.), *Human abilities: An information-processing approach*. San Francisco: Freeman.
- Kanfer, R., & Ackerman, P. L. (1989). Motivational and cognitive abilities: An integrative aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology Monograph*, *74*, 657-690.
- Kanfer, R., & Ackerman, P. L. (1990). *Ability and metacognitive determinants of skill acquisition and transfer* (Final Report 1/1/89 to 4/30/90). Bolling AFB, WA: Air Force Office of Scientific Research, Life Sciences Division (AFOSR/NL).
- Kyllonen, P. C. (1985). *Dimensions of information processing speed* (Tech. Paper AFHRL-TP-84-56). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Kyllonen, P. C. (1986). Theory-based cognitive assessment. In J. Zeidner (Ed.), *Human productivity enhancement; Organizations personnel, and decision making* (Vol. 1, pp. 338-381). New York: Praeger.
- Kyllonen, P. C. (1990, April). A taxonomy of cognitive abilities. In R. Ekstrom & I. Bejar (chairs), *New directions for the Kit of Cognitive Reference Tests*. Symposium conducted at American Educational Research Association Annual Convention, Boston, MA.
- Kyllonen, P. C. (1991). Principles for creating a computerized test battery. *Intelligence*, *15*, 1-15.
- Kyllonen, P. C. & Christal, R. E. (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. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity?! *Intelligence*, *14*, 389-433.
- Kyllonen, P. C., & Shute, V. J. (1989). A taxonomy of learning skills. In P. L. Ackerman, R. J. Sternberg, & R. Glaser (Eds.), *Learning and individual differences*. New York: Freeman.
- Kyllonen, P. C., & Stephens, D. L. (1990). Cognitive abilities as determinants of success in acquiring logic skill. *Learning and Individual Differences*, *2*, 129-160.
- Kyllonen, P. C., & Tirre, W. T. (1988). Individual differences in associative learning and forgetting. *Intelligence*, *12*, 393-421.

- Kyllonen, P. C., Tirre, W. C., & Christal, R. E. (1991). Knowledge and processing speed as determinants of associative learning proficiency. *Journal of Experimental Psychology; General*, *120*, 1–23.
- Kyllonen, P. C., & Woltz, D. J. (1989). Role of cognitive factors in the acquisition of cognitive skill. In P. Ackerman, R. Kanfer, & R. Cudeck (Eds.), *Learning and individual differences; Abilities, motivation, and methodology*, Hillsdale, NJ: Erlbaum.
- Langley, P. (1983). Exploring the space of cognitive architectures. *Behavior Research Methods and Instruments*, *15*, 289–299.
- Langley, P. (1985). Learning to search: From weak methods to domain-specific heuristics. *Cognitive Science*, *9*, 217–260.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.
- Lohman, D. F. (1979). *Spatial ability: A review and reanalysis of the correlational literature* (Tech. Rep. No. 8). Stanford, CA: Aptitude Research Project, Stanford University, School of Education.
- Lohman, D. F. (1989). Estimating individual differences in information processing using speed-accuracy models. In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology; The Minnesota symposium on learning and individual differences*, Hillsdale, NJ: Erlbaum.
- Lohman, D. F. (1991). *Estimating rate of decay of activation in semantic memory using a speed-accuracy methodology* (Final Report, No. AFOSR-89-0385). Iowa City, IA: Lindquist Center for Measurement. The University of Iowa.
- McNemar, Q. (1964). Lost: Our intelligence? Why? *American Psychologist*, *19*, 871–882.
- Metcalfe, J. (1986). Premonitions of insight predict impending error. *Journal of Experimental Psychology; Learning, Memory, and Cognition*, *12*, 623–634.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.
- Newell, A., & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Earlbaum.
- Norman, D. A. (1981). Categorization of action slips. *Psychological Review*, *88*, 1–15.
- Palmer, J., MacLeod, C. M., Hunt, E., & Davidson, J. E. (1985). Information processing correlates of reading. *Journal of Memory and Language*, *24*, 59–88.
- Pellegrino, J. W., Hunt, E. B., & Yee, P. (1989). Assessment and modeling of information coordination abilities. In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology; The Minnesota symposium on learning and individual differences*. Hillsdale, NJ: Erlbaum.
- Pena, M. C., & Tirre, W. C. (1990, April). *Cognitive correlates of programming skill acquisition*. Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA.
- Pena, M. C., & Tirre, W. C. (1990). *Cognitive factors involved in the first stage of programming skill acquisition* (AFHRL Tech Paper). Brooks AFB, TX: Air Force Human Resources Laboratory.
- Pieters, J. O. P. M., & van der Ven, Ad. H. G. S. (1982). Precision, speed, and distraction in time-limit tests. *Applied Psychological Measurement*, *6*, 93–109.
- Posner, M. (1978). *Chronometric explorations of mind*, Hillsdale, NJ: Erlbaum.

- Reason, J. T. (1984). Lapses of attention in everyday life. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention*. Orlando, FL: Academic.
- Reason, J., & Lucas, D. (1984). Absent-mindedness in shops: Its incidence, correlates, and consequences. *British Journal of Clinical Psychology*, *23*, 121–131.
- Reason, J., & Mycielska, K. (1982). *Absent-minded? The psychology of mental lapses and everyday errors*. Englewood Cliffs, NJ: Prentice-Hall.
- Ree, M. J., & Earles, J. A. (1990, May). *Differential validity of a differential aptitude test (AFHRL-TR-89-59)*. Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Revelle, W. (1989). Personality, motivation, and cognitive performance. In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology; The Minnesota symposium on learning and individual differences* (pp. 297–342). Hillsdale, NJ: Erlbaum.
- Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology; Learning, Memory, and Cognition*, *13*, 501–518.
- Shuell, T. J., & Keppel, G., (1970). Learning ability and retention. *Journal of Educational Psychology*, *61*, 59–65.
- Shute, V. J. (in press). Who is likely to acquire programming skills? *Journal of Educational Computing Research*, *7*, 1–24.
- Shute, V. J. (1991, April 1). *Individual differences research and the relationship to ITS design*. Presented at the Annual Meeting of the American Research Association, Chicago, IL.
- Shute, V. J., & Kyllonen, P. C. (1990). *Modeling individual differences in programming skill acquisition* (Tech. Rep. No. AFHRL-TP-90-76). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Snow, R. E., Chastain, R. L., & Jackson, D. (1992). *Individual differences in adaptive processing in complex learning and cognitive performance* (Final Report, Tech. Rep. No. AFOSR TR-92 0260; AD A2485860XSP). Stanford, CA: Stanford University, School of Education.
- Snow, R. E., Kyllonen, P. C., & 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: Erlbaum.
- Soule, C. S. (1990). *A cognitive skills analysis of a graph reading task*. Unpublished doctoral dissertation, University of Georgia, Athens, GA.
- Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.
- Spearman, C. (1927). *The abilities of man*. London: Macmillan.
- Squire, L. R. (1986). Mechanisms of memory. *Science*, *232*, 1612–1619.
- Sternberg, R. J. (1986). *Intelligence applied: Understanding and increasing your intellectual skills*, San Diego: Harcourt, Brace, Jovanovich.
- Sticht, T. G. (1972). Learning by listening. In R. O. Freedle & J. B. Carroll (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, DC: Winston.
- Thorndike, R. L. (1947). *Personnel selection: Test and measurement techniques*. New York: Wiley.
- Thurstone, L. L., (1938) *Primary mental abilities* (Psychometric Monograph, No. 1). Chicago, IL: University of Chicago Press.

- Thurstone, L. L., & Thurstone, T. R. (1941). *Factorial studies of intelligence* (Psychometric Monograph No. 2). Chicago, IL: University of Chicago Press.
- Tirre, W. C. (1990). *Memory and processing speed components of reading ability* (unpublished manuscript). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.
- Tirre, W. C. (in press). Knowledge and ability factors underlying simple learning by accretion. *Genetic, Social, and General Psychology Monographs*.
- Tirre, W. C., & Pena, M. C. (1992). Investigation of functional working memory in the reading span test. *Journal of Educational Psychology, 84*, 462–472.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*, 127–154.
- Underwood, B. J. (1954). Speed of learning and amount retained: A consideration of methodology. *Psychological Bulletin, 51*, 276–282.
- Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *American Psychologist, 30*, 128–134.
- Underwood, B. J., Boruch, R. F., & Malmi, R. A. (1978). Composition of episodic memory. *Journal of Experimental Psychology: General, 107*, 393–419.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The Panas scale. *Journal of Personality and Social Psychology, 54*, 1063–1070.
- Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and performance* (Vol. VIII). Hillsdale, NJ: Erlbaum.
- Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties-of-attention* (pp. 63–102). Orlando, FL: Academic.
- Woltz, D. J. (1988). An investigation of the role of working memory in procedural skill acquisition. *Journal of Experimental Psychology: General, 117*, 319–331.
- Woltz, D. J. (1990a). Decay of repetition priming effects and its relation to retention from text processing: A study of forgetting [Special Issue: Frontiers of learning and individual differences: Paradigms in transition]. *Learning and Individual Differences, 2*, 241–261.
- Woltz, D. J. (1990b). Repetition of semantic comparisons: Temporary and persistent priming effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 392–403.
- Woltz, D. J., & Shute, V. J. (1993). Individual differences in repetition priming and its relationship to declarative knowledge acquisition. *Intelligence, 17*, 333–360.
- Wothke, W., Bock, R. D., Curran, L. T., Fairbank, B. A., Augustin, J. W., Gillet, A. H., & Guerrero, C., Jr. (1990, May). *A factor analytic examination of the Armed Services Vocational Aptitude Battery (ASVAB) and the Kit of Factor-Referenced Tests* (Tech. Rep. F41689-87-D-0012/5008). Dayton, OH: Universal Energy Systems, Inc.
- Yee, P. L., Hunt, E., & Pellegrino, J. W. (1991). Coordinating cognitive information: Task effects and individual differences in integrating information from several sources. *Cognitive Psychology, 32*, 615–680.

Chapter 13

***Human Intelligence: Its Nature,
Use, and Interaction with
Context****

Robert J. Sternberg

Yale University

If conventional intelligence tests can predict only 5%–10% of the variation in various measures of life adjustment and success (see Sternberg & Wagner, 1986; Wigdor & Garner, 1982), what happened to the other 90%–95%? This chapter is an attempt to answer this question.

The main thesis of the chapter is that conventional tests are only poorly predictive of life adjustment success for at least three reasons (doubtless among others). First, the theories of intelligence on which the conventional tests are

*The work reported herein was supported under the Javits Act Program (Grant No. R206R00001) as administered by the Office of Educational Research and Improvement, U. S. Department of Education. The findings and opinions expressed in this report do not reflect the position or policies of the Office of Educational Research and Improvement or the U. S. Department of Education. Preparation of this chapter was also supported by a grant from the McDonnell Foundation.

based are too narrow. To predict beyond school, we need to understand intelligence more broadly. Second, better prediction would require us to take into account not only the levels of various aspects of intelligence, but also how these aspects of intelligence are brought to bear on the world at large through styles of thought. Two people could have very similar profiles of intelligence, but different ways of exploiting these profiles. Third, we need to take into account the kinds of contexts in which intelligence is used: Some contexts are more favorable to the exploitation of intelligence than others, or at least may be differentially favorable to different uses of it.

In this chapter, I deal with each of these three issues, first discussing the nature of intelligence; second, styles for exploiting it; and third, the kinds of contexts with which it interacts. The argument here is that intelligence can be fully understood only if we view it broadly, understand how it is exploited, and understand how it interacts with the environment.

THREE PROFILES OF INTELLIGENCE

Alice was the admissions officer's dream. She was easily admitted to the graduate program at Yale. She came with stellar test scores, and a nearly perfect record. Alice proved to be, more or less, what her record promised. She had excellent critical and analytical abilities, which helped her earn outstanding grades in her course work during the first two years at Yale. When it came to taking tests and writing course term papers, she had no peer among her classmates. During her first couple of years in the graduate program, she was an outstanding success. But after the first two years, Alice no longer looked quite so outstanding. In our graduate program, as in most, emphasis shifts after the first couple of years. It is not enough just to criticize other people's ideas, or to study concepts that other people have proposed. You must start coming up with your own ideas, and figuring out ways of implementing them. Alice's synthetic abilities were far inferior to her analytic ones, but there would have been no way of knowing this fact from the evidence available in the admissions folder. For although conventional measures can give us a good reading on analytic abilities, they give virtually no reading on synthetic abilities. Thus, Alice was "IQ-test" smart, but she was not equally intelligent in the synthetic or practical areas of intelligence.

People like Barbara are the admissions officer's nightmare. When she applied to Yale, she had good grades, but abysmal aptitude test scores, at least by Yale standards. Despite these low scores, she had superlative letters of recommendation, which described her as an exceptionally creative young woman, who had designed and implemented creative research with only the most minimal guidance. Moreover, her resume showed her to have been actively

involved in important research. Unfortunately, people like Barbara are rejected from many graduate programs. As a result, they either have to enter a program that is much less competitive or enter a different field altogether.

This pattern of events is not limited to graduate school. There are thousands of people like Barbara who are rejected in a similar way from law schools, medical schools, business schools, education schools, and the like. Some of them never even get to this point, having been rejected earlier from competitive colleges. However, sometimes there are exceptions and in these instances people like Barbara often show themselves to be fine students possessing excellent research abilities. These people may not excel in course performance, although they might do much better than test scores predict. But when the demands of the graduate program shift, for example, to an emphasis on synthetic abilities, people like Barbara are in their element. They may not have Alice's analytic abilities, but they greatly surpass Alice in synthetic abilities.

Celia, on paper, appeared to be somewhere between Alice and Barbara in terms of suitability for admission to the graduate program. She was good on almost every measure of success, but not truly outstanding on any of them. We admitted her, expecting her to come out near the middle of the class. This did not happen. Celia has proved to be outstanding, although in a way that is quite different from Alice or Barbara. Celia's expertise is in figuring out and in adapting to the demands of the environment. Placed in a new kind of setting, she can figure out what is required of her, and then go ahead and do it just right. She knows exactly what to do to get ahead. In conventional parlance, Celia is "street-smart." She excels in practical intelligence.

Just how might one characterize the similarities and differences among Alice, Barbara, and Celia? Clearly, all of them are exceedingly intelligent, although in very different ways. People like Alice excel in conventional academic or analytic intelligence. To the extent that one seeks to understand intelligence in terms of the conventional factors or information-processing components that researchers have used to characterize intelligence, individuals such as Alice would be viewed as very, very smart. Thus, if one looks at the relationship between intelligence and the internal world of the individual, people like Alice excel. Individuals like Barbara do not look nearly so intelligent in terms of conventional notions of academic intelligence. Where they excel is in synthetic ability, or the ability to deal with novelty—to view new things in old ways or old things in new ways. Hence, these people come out looking extremely intelligent if one looks at the relationships of intelligence to experience, and particularly, novel experience. People like Celia have neither Alice's nor Barbara's patterns of strength. Instead, they excel in terms of the relationship between intelligence and the external world of the individual. Their excellence is in practical intelligence, or in applying their mental abilities to everyday kinds of situations. Their street smarts are not measured by conventional tests, but quickly show up in their performance in real-world settings.

THE TRIARCHIC THEORY OF HUMAN INTELLIGENCE

The triarchic theory of human intelligence seeks to explain in an integrative way the relationship between (a) intelligence and the internal world of the individual, or the mental mechanisms that underlie intelligent behavior; (b) intelligence and the external world of the individual, or the use of these mental mechanisms in everyday life in order to attain an intelligent fit to the environment; and (c) intelligence and experience, or the mediating role of one's passage through life between the internal and external worlds of the individual. Consider some of the basic tenets of the theory.

Intelligence and the Internal World of the Individual

Psychometricians, Piagetians, and information-processing psychologists have all recognized the importance of understanding what mental states or processes underlie intelligent thought. In the triarchic theory, this understanding is sought through the identification and understanding of three basic kinds of information-processing components, which are referred to as metacomponents, performance components, and knowledge-acquisition components.

Metacomponents. Metacomponents are *higher order*, executive processes used to plan what one is going to do, to monitor it while one is doing it, and to evaluate it after it is done. These metacomponents include (a) recognizing the existence of a problem, (b) deciding on the nature of the problem confronting one, (c) selecting a set of lower order processes to solve the problem, (d) selecting a strategy into which to combine these components, (e) selecting a mental representation upon which the components and strategy can act, (f) allocating one's mental resources, (g) monitoring one's problem solving as it is happening, and (h) evaluating one's problem solving after it is done. Consider some examples of some of these higher order processes.

Deciding upon the nature of a problem plays a prominent role in intelligence. For example, with young children as well as older adults, difficulty in problem solving often lies not in actually solving a given problem, but in figuring out just what the problem is that needs to be solved (see, for example, Flavell, 1977; Sternberg & Rifkin, 1979). A major feature distinguishing retarded persons from normal ones is the retardates' need to be instructed explicitly and completely as to the nature of the particular task they are solving and how it should be performed (Butterfield, Wambold, & Belmont, 1973; Campione & Brown, 1979). The importance of figuring out the nature of the problem is not limited to retarded persons. Resnick and Glaser (1976) have argued that intelligence is the ability to learn from incomplete instruction.

Selection of a strategy for combining lower order components is also a critical aspect of intelligence. In early information-processing research on intelli-

gence, including my own (e.g., Sternberg, 1977), the primary emphasis was simply on figuring out what subjects do when confronted with a problem. What components do subjects use, and into what strategies do they combine these components? Soon, however, information-processing researchers began to ask the question of why subjects use the strategies they choose. For example, Cooper (1982) has reported that in solving spatial problems, and especially mental-rotation problems, some subjects seem to use a holistic strategy of comparison whereas others use an analytic strategy. She has sought to figure out what leads subjects to the choice of one strategy or another. Siegler (1986) has actually proposed a model of strategy selection in arithmetic computation problems that links strategy choice to both the rules and mental associations one has stored in long-term memory. MacLeod, Hunt, and Mathews (1978) found that high-spatial subjects tend to use a spatial strategy in solving sentence-picture comparison problems, whereas high-verbal subjects are more likely to use a linguistic strategy. In my own work, I have found that subjects tend to prefer certain strategies for analogical reasoning over others because they place fewer demands upon working memory (Sternberg & Ketron, 1982). Similarly, subjects choose different strategies in linear-syllogistic reasoning (spatial, linguistic, mixed spatial-linguistic), but in this task, they do not always capitalize on their ability patterns so as to choose the strategy most suitable to their respective levels of spatial and verbal abilities (Sternberg & Weil, 1980). In sum, the selection of a strategy seems to be at least as important for understanding intelligent task performance as is the efficacy with which the chosen strategy is implemented.

Intimately tied up with the selection of a strategy is the selection of a mental representation for information. In the early literature on mental representations, the emphasis seemed to be on understanding how information is represented. For example, can individuals use imagery as a form of mental representation (Kosslyn, 1980)? In more recent research, investigators have realized that people are quite flexible in their representations of information. The most appropriate question to ask seems to be not how is information represented, but which representations are used in what circumstances? For example, Sternberg (1977) found that analogy problems using animal names can draw on either spatial or clustering representations of the animal names. In the studies of strategy choice mentioned earlier, it was found that subjects can use either linguistic or spatial representations in solving sentence-picture comparisons (MacLeod et al., 1978) or linear syllogisms (Sternberg & Weil, 1980). Sternberg and Rifkin (1979) found that the mental representation of certain kinds of analogies can be either more or less holistic, depending upon the age of the subjects.

As important as any other metacomponent is one's ability to allocate one's mental resources. Different investigators have studied resource allocation in different ways. Hunt and Lansman (1982), for example, have concentrated on

the use of secondary tasks in assessing information processing, and have proposed a model of attention allocation in the solution of problems that involves both a primary and a secondary task. In my work, I have found that better problem solvers tend to spend relatively more time in global strategy planning (Sternberg, 1981). Similarly, in solving analogies, better analogical reasoners seem to spend relatively more time encoding the terms of the problem than do poorer reasoners, but to spend relatively less time in operating on these encodings (Sternberg, 1977; Sternberg & Rifkin, 1979). In reading as well, the better readers are better able than the poorer readers to allocate their time across reading passages as a function of the difficulty of the passages to be read, and of the purpose for which the passages are being read (cf. Brown, Bransford, Ferraro, & Campione, 1983; Wagner & Sternberg, 1987).

Finally, monitoring one's solution processes is a key aspect of intelligence (see also Brown, 1978). Consider, for example, the Missionaries and Cannibals problem, in which the subjects must "transport" a set of missionaries and cannibals across a river in a small boat without allowing the cannibals an opportunity to eat the missionaries, an event that can transpire only if the cannibals are allowed to outnumber the missionaries on either side of the river bank. The main kinds of errors that can be made are either to return to an earlier state in the problem space for solution, or to make an impermissible move (Simon & Reed, 1976; see also Sternberg, 1982b). Neither of these errors would result if a given subject closely monitored his or her solution processes. For young children learning to count, a major source of errors in counting objects is to count a given object twice, an error that, again, can result from a failure in solution monitoring (Gelman & Gallistel, 1978). The effects of solution monitoring are not limited, of course, to any one kind of problem. One's ability to use the strategy of means-ends analysis (Newell & Simon, 1972)—that is, reduction of differences between where one is in solving a problem and where one wishes to get in solving that problem—depends on one's ability to monitor just where one is in problem solution.

Performance components. Performance components are *lower order* processes that execute the instructions of the metacomponents. These lower order components solve the problems according to the plans laid out by the metacomponents. Whereas the number of metacomponents used in the performance of various tasks is relatively limited, the number of performance components is probably quite large. Many of these performance components are relatively specific to narrow ranges of tasks (Sternberg, 1979, 1983, 1985a).

One of the most interesting classes of performance components is that found in inductive reasoning of the kind measured by tests such as matrices, analogies, series completions, and classifications. These components are important because of the importance of the tasks into which they enter: Induction problems of these kinds show the highest loadings on the so-called *g*, or general intelligence factor (Jensen, 1980; Snow & Lohman, 1984; Sternberg &

Gardner, 1982). Thus, identifying these performance components can give us some insight into the nature of the general factor. In saying this, I am not arguing for any one factorial model of intelligence (i.e., one with a general factor) over others: To the contrary, I believe that most factor models are mutually compatible, differing only in the form of rotation that has been applied to a given factor space (Sternberg, 1977). The rotation one uses is a matter of theoretical or practical convenience, not of truth or falsity.

The main performance components of inductive reasoning are encoding, inference, mapping, application, comparison, justification, and response. They can be illustrated with reference to an analogy problem, such as LAWYER : CLIENT :: DOCTOR : (a) PATIENT, (b) MEDICINE. In encoding, the subject retrieves from semantic memory semantic attributes that are potentially relevant for analogy solution. In inference, the subject discovers the relation between the first two terms of the analogy, here, LAWYER and CLIENT. In mapping, the subject discovers the higher order relation that links the first half of the analogy, headed by LAWYER, to the second half of the analogy, headed by DOCTOR. In application, the subject carries over the relation inferred in the first half of the analogy to the second half of the analogy, generating a possible completion for the analogy. In comparison, the subject compares each of the answer options to the mentally generated completion, deciding which, if any, is correct. In justification, used optionally if none of the answer options matches the mentally generated solution, the subject decides which, if any, of the options is close enough to constitute an acceptable solution to the examiner, whether by means of pressing a button, making a mark on a piece of paper, or whatever.

Two fundamental issues have arisen regarding the nature of performance components as a fundamental construct in human intelligence. The first, mentioned briefly above, is whether their number simply keeps expanding indefinitely. Neisser (1983), for example, has suggested that it does. As a result, he views the construct as of little use. But this expansion results only if one considers seriously those components that are specific to small classes of problems or to single problems. If one limits one's attention to the more important, general components of performance, the problem simply does not arise, as shown, for example, in Sternberg and Gardner's (1982) analysis of inductive reasoning, or in Pellegrino and Kail's (1982) analysis of spatial ability. The second issue is one of the level at which performance components should be studied. In so-called "cognitive correlates" research (Pellegrino & Glaser, 1979), theorists emphasize components at relatively low levels of information processing (Hunt, 1978, 1980; Jensen, 1982). In so-called "cognitive components" research (Pellegrino & Glaser, 1979), theorists emphasize components at relatively high levels of information processing (e.g., Mulholland, Pellegrino, & Glaser, 1980; Snow, 1979; Sternberg, 1977). Because of the interactive nature of human information processing, it would appear that there is no right or

wrong level of analysis. Rather, all levels of information processing contribute to both task and subject variance in intelligent performance. The most expeditious level of analysis depends upon the task and subject population: Lower level performance components might be more important, for example, in studying more basic information-processing tasks, such as choice reaction time, or in studying higher level tasks, but in children who have not yet automatized the lower processes that contribute to performance on these tasks.

Knowledge-acquisition components. Knowledge-acquisition components are used to *learn how to do* what the metacomponents and performance components eventually do. Three knowledge-acquisition components appear to be central in intellectual functioning: (a) selective encoding, (b) selective combination, and (c) selective comparison.

Selective encoding involves sifting out relevant from irrelevant information. When new information is presented in natural contexts, relevant information for one's given purpose is embedded in the midst of large amounts of purpose-irrelevant information. A critical task for the learner is that of sifting the "wheat from the chaff": recognizing just what information among all the pieces of information is relevant for one's purposes (see Schank, 1980).

Selective combination involves combining selectively encoded information in such a way as to form an integrated, plausible whole. Simply sifting out relevant from irrelevant information is not enough to generate a new knowledge structure. One must know how to combine the pieces of information into an internally connected whole (see Mayer & Greeno, 1972).

My emphasis on components of knowledge acquisition differs somewhat from the focus of some contemporary theorists in cognitive psychology, who emphasize what is already known, and the structure of this knowledge (e.g., Chase & Simon, 1973; Chi, 1978; Keil, 1984). I should point out, again, therefore, that these various emphases are complementary. If one is interested in understanding, for example, differences in performance between experts and novices, clearly one would wish to look at the amount and structure of their respective knowledge bases. But if one wishes to understand how these differences came to be, merely looking at developed knowledge would not be enough. Rather, one would have to look as well at differences in the ways in which the knowledge bases were acquired. It is here that understanding of knowledge-acquisition components will prove to be most relevant.

We have studied knowledge-acquisition components in the domain of vocabulary acquisition (e.g., Sternberg, 1987; Sternberg & Powell, 1983). Difficulty in learning new words can be traced, at least in part, to the application of components of knowledge acquisition to context cues stored in long-term memory. Individuals with higher vocabularies tend to be those who are better able to apply the knowledge-acquisition components to vocabulary-learning situations. Given the importance of vocabulary for overall intelligence,

almost without respect to the theory or test one uses, utilization of knowledge-acquisition components in vocabulary-learning situations would appear to be critically important for the development of intelligence. Effective use of knowledge-acquisition components is trainable. I have found, for example, that just 45 minutes of training in the use of these components in vocabulary learning can significantly and fairly substantially improve the ability of adults to learn vocabulary from natural-language contexts (Sternberg, 1987).

To summarize, then, the components of intelligence are an important part of the intelligence of the individual. The various kinds of components work together. Metacomponents activate performance and knowledge-acquisition components. These latter kinds of components in turn provide feedback to the metacomponents. Although one can isolate various kinds of information-processing components from task performance using experimental means, in practice, the components function together in highly interactive, and not easily isolable, ways. Thus, diagnoses as well as instructional interventions need to consider all three types of components in interaction, rather than any one kind of component in isolation. But understanding the nature of the components of intelligence is not, in itself, sufficient to understand the nature of intelligence, because there is more to intelligence than a set of information-processing components. One could scarcely understand all of what it is that makes one person more intelligent than another by understanding the components of processing on, say, an intelligence test. The other aspects of the triarchic theory address some of the other aspects of intelligence that contribute to individual differences in observed performance, outside of testing situations as well as within them.

Intelligence and Experience

Components of information processing are always applied to tasks with which one has some level of prior experience (including the null level) and in situations with which one has some level of prior experience (including the null level). Hence, these internal mechanisms are closely tied to one's experience. According to the experiential subtheory, the components are not equally good measures of intelligence at all levels of experience. Assessing intelligence requires one to consider not only components, but the level of experience at which they are applied.

During recent years, there has been a tendency in cognitive science to study script-based behavior (e.g., Schank & Abelson, 1977), whether under the name of "script" or under some other name, such as "schema" or "frame." There is no longer any question that much of our behavior is scripted, in some sense. However, from the standpoint of the present subtheory, such behavior is nonoptimal for understanding intelligence. Typically, one's actions when one

goes to a restaurant or a doctor's office or a movie theater do not provide good measures of intelligence, even though they do provide good measures of scripted behavior. What, then, is the relation between intelligence and experience?

According to the experiential subtheory, intelligence is best measured at those regions of the experiential continuum that involve tasks or situations that are either relatively novel, on the one hand, or in the process of becoming automatized, on the other. As Raaheim (1974) pointed out, totally novel tasks and situations provide poor measures of intelligence: One would not want to administer, say, trigonometry problems to a first-grader of roughly 6 years of age. But one might wish to administer problems that are just at the limits of a child's understanding, in order to test how far this understanding extends. Related is Vygotsky's (1978) concept of the zone of proximal development, in which one examines a child's ability to profit from instruction to facilitate his or her solution of novel problems. In order to measure automatization skill, one might wish to present a series of problems—mathematical or otherwise—and to see how long it takes for solution of them to become automatic, and to see how automatized performance becomes. Thus, both slope and asymptote (if any) of automatization are of interest.

Ability to deal with novelty. Several sources of evidence converge upon the notion that the ability to deal with relative novelty is a good way of measuring intelligence. Consider three such sources of evidence. First, we have conducted several studies on the nature of insight, both in children and in adults (Davidson & Sternberg, 1984; Sternberg & Davidson, 1982). In the studies with children (Davidson & Sternberg, 1984), we separated three kinds of insights: insights of selective encoding, insights of selective combination, and insights of selective comparison. Use of these knowledge-acquisition components is referred to as insightful when they are applied in the absence of existing scripts, plans, frames, or whatever. In other words, one must decide what information is relevant, decide how to put the information together, or decide how new information relates to old in the absence of any obvious cues on the basis of which to make these judgments. A problem is insightfully solved at the individual level when a given individual lacks such cues. A problem is insightfully solved at the societal level when no one else has these cues either. In these studies, our hypothesis was that children who are intellectually gifted are gifted in part by virtue of their insight abilities, which represent an important part of the ability to deal with novelty.

Children were administered quantitative insight problems, of the kinds found in puzzle books, that measured primarily either selective encoding skill, selective combination skill, or selective comparison skill. We manipulated the need for such insights experimentally. Problems were either administered uncued (standard format) or precued. The form of precueing depended upon the kind of insight being assessed. For selective encoding, we precued what information was relevant for solving a given problem by highlighting all infor-

mation that was relevant for solving each problem. Thus, we eliminated the need for selective encoding by pointing out to the children just what information was relevant for each given problem. In the selective-combination condition, precueing consisted of information telling the children how to combine the given information selectively. For example, a table might be drawn showing how the various terms of the problem interrelated. In the selective-comparison condition, the need for selective comparison was manipulated by varying the examples in the introduction to the problems. Precued conditions ranged from one in which examples were given but their relevance to the later problems not pointed out, to examples that were explicitly stated to be relevant for solution of the later problems, to examples that were indicated as relevant to designated problems in the set of problems that needed to be solved. The basic design, therefore, was to test children either identified as gifted or not so identified, and to administer problems that either required insights of one of the three kinds or that did not require such insights because the insights were provided to the children.

The critical finding was that providing insights to the children significantly benefited the nongifted, but not the gifted children. (None of the children performed anywhere near ceiling, so that the interaction was not due to ceiling effects.) In other words, the gifted children spontaneously had the insights and hence did not benefit from being given these insights. The nongifted children did not have the insights spontaneously, and hence did benefit. Thus, the gifted children were better able spontaneously to deal with novelty.

In a very different paradigm, adult subjects were given what I call conceptual-projection problems (Sternberg, 1982a). In these problems, one has to make predictions about future states of objects based upon incomplete and sometimes partially faulty information about the current states of the objects. These problems generally employed a science-fiction type of scenario. For example, one might be introduced to four kinds of people on the planet Kyrion: One kind of person is *born young and dies young*, a second kind of person is *born young and dies old*, a third kind is *born old and dies old*, and a fourth kind is *born old and dies young*. Given incomplete information about the person in the present, one has to figure out what kind of person the individual is (names such as "kwef," "pros," "balt," and "plin" were used) and determine what his or her appearance would be 20 years later. Performance on the conceptual-projection task was experimentally decomposed, and the mathematical model of task performance accounted for most of the stimulus variance (generally 90+%) in task performance.

Each of these component scores was then correlated with performance on a variety of psychometric tests, including tests of inductive reasoning ability, which are primary measures of general intelligence. The critical finding was that the correlation of overall response time (generally at the level of about -.6, negative because response times were correlated with numbers correct) with

psychometric test scores was due to correlations stemming from those performance components tapping the ability to deal with novelty, for example, changing conceptual systems from a familiar one (born young and dies old) to an unfamiliar one (e.g., born old and dies young). These correlations held up without regard to the particular surface structure of the problem, of which the scenario about birth and death states was only one of four. Thus, it was the ability to deal with novelty, rather than other abilities involved in solving the problems, that proved to be critical to general intelligence.

A third source of evidence for the proposed hypothesis derives from the large literature on fluid intelligence, which is in part a kind of intelligence that involves dealing with novelty (see Cattell, 1971). Snow and Lohman (1984; see also Snow, Kyllonen, & Marshalek, 1984) have multidimensionally scaled a variety of such tests and found the dimensional loadings to follow a radex structure. In particular, tests with higher loadings on *g*, or general intelligence, fall closer to the center of the spatial diagram. The critical thing to note is that those tests that best measure the ability to deal with novelty fall closer to the center, and tests tend to be greater removed from the center as their assessment of the ability to deal with novelty becomes more remote. In sum, evidence from the laboratories of others as well as myself supports the idea with the various components of intelligence that are involved in dealing with novelty, as measured in particular tasks and situations, provide particularly apt measures of intellectual ability.

Ability to automatize information processing. Although we are only now testing the second aspect of the experiential subtheory, that of the ability to automatize information processing, there are several converging lines of evidence in the literature to support the claim that this ability is a key aspect of intelligence. For example, Sternberg (1977) found that the correlation between people-piece (schematic-picture) analogy performance and measures of general intelligence increased with practice, as performance on these items became increasingly automatized. Skilled reading is heavily dependent upon automatization of bottom-up functions, and the ability to read well is an essential part of crystallized ability, whether as viewed from the standpoint of theories such as Cattell's (1971) or Vernon's (1971), or from the standpoint of tests of crystallized ability, such as the verbal portion of the Scholastic Aptitude Test. Poor comprehenders often are those who have not automatized the elementary, bottom-up processes of reading, and hence who do not have sufficient attentional resources to allocate to top-down comprehension processes.

Theorists such as Jensen (1982) and Hunt (1978) have attributed the correlation between tasks such as choice reaction time and letter matching to the relation between speed of information processing and intelligence. Indeed, there is almost certainly some relation, although I believe it is much more complex than these theorists seem to allow for. But a plausible alternative hypothesis is that at least some of that correlation is due to the effects of

automatization of processing: Because of the simplicity of these tasks, they probably become at least partially automatized fairly rapidly, and hence can measure both rate and asymptote of automatization of performance. In sum, then, although the evidence is far from complete, there is at least some support for the notion that rate and level of automatization are related to intellectual skill.

The ability to deal with novelty and the ability to automatize information processing are interrelated, as shown in the example of reading above. If one is well able to automatize, one has more resources left over for dealing with novelty. Similarly, if one is well able to deal with novelty, one has more resources left over for automatization. Thus, performance at the various levels of the experiential continuum are related to one another.

These abilities should not be viewed in a vacuum with respect to the componential subtheory. The components of intelligence are applied to tasks and situations at various levels of experience: The ability to deal with novelty can be understood in part in terms of the metacomponents, performance components, and knowledge-acquisition components involved in it. Automatization, when it occurs, is of these components. Hence, the two subtheories considered so far are closely intertwined. We need now to consider the application of these subtheories to everyday tasks, in addition to laboratory ones.

Intelligence and the External World of the Individual

According to the contextual subtheory, intelligent thought is directed toward one or more of three behavioral goals: *adaptation to an environment*, *shaping of an environment*, or *selection of an environment*. These three goals may be viewed as the functions toward which intelligence is directed: Intelligence is not aimless or random mental activity that happens to involve certain components of information processing at certain levels of experience. Rather, it is purposefully directed toward the pursuit of these three global goals, all of which have more specific and concrete instantiations in people's lives.

Adaptation. Most intelligent thought is directed toward the attempt to adapt to one's environment. The requirements for adaptation can differ radically from one environment to another—whether environments are defined in terms of families, jobs, subcultures, cultures, or whatever. Hence, although the components of intelligence required in these various contexts may be the same or quite similar, and although all of them may involve, at one time or another, dealing with novelty and automatization of information processing, the concrete instantiations that these processes and levels of experience take may differ substantially across contexts. This fact has an important implication for our understanding of the nature of intelligence. According to the triarchic theory, in general, and the contextual subtheory, in particular, the processes and experi-

ential facets and functions of intelligence remain essentially the same across contexts; but the particular instantiations of these processes, facets, and functions can differ, and differ radically. Thus, the content of intelligent thought and its manifestations in behavior will bear no necessary resemblance across contexts. As a result, although the mental elements that an intelligence test should measure do not differ across contexts, the vehicle for measurement may have to differ. A test that measures a set of processes, experiential facets, or intelligent functions in one context may not provide equally adequate measurement in another context. To the contrary, what is intelligent in one culture may be viewed as unintelligent in another.

A nice example of this fact can be found in the work of Cole, Gay, Glick, and Sharp (1971). These investigators asked adult Kpelle tribesmen to sort 20 familiar objects into groups of things that belong together. Their subjects separated the objects into functional groupings (e.g., a knife with an orange), as children in Western societies would do. This pattern of sorting surprised the investigators, who had expected to see taxonomic groupings (e.g., tools sorted together and foods sorted together) of the kind that would be found in the sortings of Western adults. Had the investigators used the sorting task as a measure of intelligence in the traditional way, they might well have labeled the Kpelle tribesmen as intellectually inferior to Western adults. However, through persistent exploration of why the Kpelle were sorting in this way, they found that the Kpelle considered functional sorting to be the intelligent form of sorting. When the tribesmen were asked to sort the way a stupid person would, they had no trouble sorting taxonomically. In short, they differed on this task not in their intellectual competence vis a vis Western adults, but in their conception of what was functionally adaptive. Indeed, it takes little thought to see the practicality of sorting functionally: People do, after all, use utensils in conjunction with foods of a given category (e.g., fruits) on a frequent basis.

In the case of Kpelle tribesmen, different contextual milieux resulted in a different conception of what constitutes intelligence: The particular difference illustrated earlier is in what is considered to be adaptive, rather than in the ability to act adaptively. But different contextual milieux may result in the development of different mental abilities. For example, Puluwat navigators must develop their large-scale spatial abilities for dealing with cognitive maps to a degree that far exceeds the adaptive requirements of contemporary Western societies (Gladwin, 1970). Similarly, Kearins (1981) found that aboriginal children probably develop their visuospatial memories to a greater degree than do Anglo-Australian children, who are more likely to apply verbal strategies to spatial memory tasks than are the aborigines, who employ spatial strategies. In contrast, participants in Western societies probably develop their abilities for thinking abstractly to a greater degree than do societies in which concepts are rarely dealt with outside their concrete manifestations in the objects of the everyday environment.

One of the most interesting differences among cultures and subcultures in the development of patterns of adaptation is in the matter of time allocation, a metacomponential function. In Western cultures, in general, budgeting of time and careful allocation of one's time to various activities is a prized commodity. Our lives are largely governed by careful scheduling at home, in school, at work, and so on. There are fixed hours for certain activities, and fixed lengths of time within which these activities are expected to be completed. Indeed, the intelligence tests we use show our prizing of time allocation to the fullest. Almost all of them are timed in such a way as to make completion of the tests a nontrivial challenge. A slow or very cautious worker is at a distinct disadvantage.

Not all cultures and subcultures view time in the same way that we do. For example, among the Kipsigi, schedules are much more flexible, and hence these individuals have difficulty understanding and dealing with Western notions of the time pressure under which people are expected to live (Super & Harkness, 1980). In Hispanic cultures, such as Venezuela, my own personal experience indicates that the press of time is taken with much less seriousness than it is in typical North American cultural settings. Even within the continental United States, though, there can be major differences in the importance of time allocation. Heath (1983) describes young children brought up in the rural community of "Trackton," in which there is very little time pressure and in which things essentially get done when they get done. These children can have great difficulty adjusting to the demands of the school, in which severe time pressures may be placed upon the children for the first time in their lives.

The point of these examples has been to illustrate how differences in environmental press and people's conceptions of what constitutes an intelligent response to it can influence just what counts as adaptive behavior. To understand intelligence, one must understand it not only in relation to its internal manifestations in terms of mental processes, and its experiential manifestations in terms of facets of the experiential continuum, but also in terms of how thought is intelligently translated into action in a variety of different contextual settings. The differences in what is considered adaptive and intelligent can extend even to different occupations within a given cultural milieu. For example, Sternberg (1985b) has found that individuals in different fields of endeavor (art, business, philosophy, physics) view intelligence in slightly different ways that reflect the demands of their respective fields.

Shaping. Shaping of the environment is often used when adaptation fails, as a backup strategy. If one is unable to change oneself so as to fit the environment, one may attempt to change the environment so as to fit oneself. For example, repeated attempts to adjust to the demands of one's romantic partner may eventually lead to attempts to get the partner to adjust to oneself. But shaping is not always used in lieu of adaptation. In some cases, shaping may be used before adaptation is ever tried, as in the case of the individual who

attempts to shape a romantic partner with little or no effort to shape him or herself so as better to suit the partner's wants or needs.

In the laboratory, examples of shaping behavior can be seen in strategy selection situations where one essentially molds the task to fit one's preferred style of dealing with tasks. For example, in comparing sentence statements to pictures that either do or do not accurately represent these statements, individuals may select either a verbal or a spatial strategy, depending upon their pattern of verbal and spatial abilities (MacLeod, Hunt, & Mathews, 1978). The task is "made over" in conformity to what one does best. Similarly, I find that in multivariate statistics, my graduate students tend to view problems either algebraically or geometrically, depending upon their pattern of abilities and preferences. My own presentation of the subject is, again, "made over" to conform to their needs and desires.

Because people operate in groups as well as in isolation, attempts by group members to shape in different ways can result in products that either profit or lose from the group effort. I have recently attended a rather unstructured meeting in which a group of individuals attempted to accomplish a variety of agendas. But because of limited resources, not all of the agendas could be realized. The result was that practically none of them were, because of attempts by individuals to realize their own agendas at the expense of other people's. A more salutary result has eventuated from my collaborations with one of my graduate students. In research, I tend to be a "selective comparer," constantly seeking to relate new theories and facts to old ones. I am probably less careful, however, about selective combination, that is, about fitting together the various facts at my disposal. As a result, I may neglect to deal with those facts that do not quite fit into the framework I establish for them. My graduate student tends to be more a selective combiner than comparer. She is less concerned with relating new facts to old facts or theories, but more concerned with making sure that the various new facts can be fitted together into a coherent account that deals with them all. In our collaborations in research, we each attempt to shape the outcomes of the research in accordance with our preferred style of working. In this case, the two styles complement each other, as one individual makes sure that the research is not conducted in isolation from past research or ideas, whereas the other individual makes sure that inconvenient experimental results are not shunted aside. Indeed, it is such results that may eventuate in the true breakthroughs in research. Thus, in this case, the attempts to shape the environment in two different ways result in a healthy tension that improves rather than harms the final outcome.

In some respects, shaping may be seen as the quintessence of intelligent thought and behavior. One essentially makes over the environment rather than allowing the environment to make over oneself. Perhaps it is this skill that has

enabled humankind to reach its current level of scientific, technological, and cultural advancement (for better or for worse). In science, the greatest scientists are those who set the paradigms (shaping), rather than those who merely follow them (adaptation). Similarly, in art and in literature, the individuals who achieve greatest distinction are often those who create new modes and styles of expression, rather than merely following existing ones. It is not their use of shaping alone that distinguishes them intellectually, but rather a combination of their willingness to do it with their skill in doing it.

Selection. Selection involves renunciation of one environment in favor of another. In terms of the rough hierarchy established so far, selection is sometimes used when both adaptation and shaping fail. After attempting both to adapt to and shape a marriage, one may decide to deal with one's failure in these activities by "deselecting" the marriage, and choosing the environment of the newly single. Failure to adjust to the demands of work environments, or to change the demands placed upon one so as to make them a reasonable fit to one's interests, values, expectations, or abilities, may result in the decision to seek another job altogether. But selection is not always used as a last resort. Sometimes one attempts to shape an environment only after attempts to leave it have failed. Other times, one may decide almost instantly that an environment is simply wrong for oneself, and feel that one need not or should not even try to fit into or to change it. For example, we get, every now and then, a new graduate student who realizes almost immediately that he or she came to graduate school for the wrong reason, or who finds that graduate school is nothing at all like the continuation of undergraduate school he or she expected. In such cases, the intelligent thing to do may be to leave the environment as soon as possible, in order to pursue activities more in line with one's goals in life.

Environmental selection is not usually directly studied in the laboratory, although it may have relevance for certain experimental settings. Perhaps no research example of its relevance has been more salient than the experimental paradigm created by Milgram (1975), who, in a long series of studies, asked subjects to "shock" other subjects (who were actually confederates and who were not shocked). The finding of critical interest was how few subjects either shaped the environment by refusing to shock their victims, or employed the device of selection by simply refusing to continue with the experiment and walking out of it. Milgram has drawn an analogy to the situation in Nazi Germany, where obedience to authority created an environment whose horrors continue to amaze us to this day, and always will. This example is a good one in showing how close matters of intelligence can come to matters of personality. In fact, many Jews refused to leave Nazi-occupied territories for fear of losing their property, their peers, and so on. Their refusal may have been due

to personality factors, but for many of them, their decision to stay was in some respects the supreme act of unintelligence, as it resulted in their death, not through choice, but later through having no choice at all in the matter.

To conclude, adaptation, shaping, and selection are functions of intelligent thought as it operates in context. They may, although they need not, be employed hierarchically, with one path followed when another one fails. It is through adaptation, shaping, and selection that the components of intelligence, as employed at various levels of experience, become actualized in the real world. In this section, it has become clear that the modes of actualization can differ widely across individuals and groups, so that intelligence cannot be understood independently of the ways in which it is manifested.

The goal of the triarchic theory is not to replace previous theories of intelligence, but rather, to incorporate them, and particularly, their best aspects. I argued earlier that most theories of intelligence are intercompatible, and this argument applies to the triarchic theory as well, *vis-à-vis* other theories of intelligence. Consider why.

The earliest psychometric theory was Galton's. Galton placed heavy stress on psychophysical measures, and many of these, such as strength of grip, are given no credence today, except perhaps in Gardner's (1983) description of multiple intelligences, according to which strength might be part of kinesthetic intelligence. However, Galton's mental measures of attributes such as reaction time were a step in the right direction. These measures probably assessed, in some degree, speed of performance-componential functioning as well as speed of performance component execution, and both of these attributes are of some importance to intelligence. If Galton's and, later, Cattell's measures were not highly correlated with anything else, one would expect such low correlations due to the very narrow aspect of intelligence that these measures assessed.

The triarchic theory is more in line with the theory of Binet and Simon (1973), with its emphasis upon judgment. Binet emphasized higher order thinking in his theory, and his tests involved substantial investment of both meta-components and performance components. Some of these tests, moreover, apply these components in novel situations. A few even apply them in contextually relevant ones. Hence, Binet's theory was more in the spirit of later theories of intelligence, including the triarchic one.

The first major factor theory of intelligence was that of Spearman (1923, 1927). Spearman's theory had two parts, the purely psychometric part, according to which intelligence is understood primarily in terms of a general factor, and an information-processing part, which specified three of the mental processes Spearman believed to be central in general intelligence. These three

processes—*apprehension of experience*, *eduction of relations*, and *eduction of correlates*—are essentially identical to the performance components of *encoding*, *inference*, and *application*, respectively, that appear in the componential theory of induction. We, too, have found these components to be of central importance in the understanding of general intelligence (Sternberg & Gardner, 1982). We believe, with Spearman, that there is more to general intelligence than these three processes, however. This something more would include meta-componential functioning as well as the execution of general components of information processing (including components from all three categories) that are common across almost all information-processing tasks.

Thurstone's (1938) theory of primary mental abilities was originally seen by many as conflicting with Spearman's theory. It has become clear in recent hierarchical theories, and especially Gustafsson's (1984), that Spearman's and Thurstone's theories may be viewed as compatible, simply because they deal with factors at different levels of a hierarchy of decreasing generality. I view Thurstone's theory as tapping primarily the *class components* that apply to certain classes or groups of tasks (e.g., spatial tasks, verbal comprehension tasks, memory tasks) but not others (Sternberg, 1980). The performance component involved in mental rotation of objects is limited to spatial tasks and reasoning tasks that involve mentally rotating figures. Thus, this theory, too, is in some sense a subset of the triarchic one.

Probably the most well-known hierarchical theories are those of Cattell (1971) and of Vernon (1971). The theories are rather similar, revealing Burt's influence, and I shall not try to deal with all aspects of either theory. In Cattell's theory, the most well-known part of it deals with two subfactors of general intelligence, *fluid ability* and *crystallized ability*. Fluid ability is best measured by tasks requiring abstract reasoning, whereas crystallized ability is best measured by tasks requiring demonstration of accumulated knowledge, such as vocabulary. From the standpoint of the triarchic theory, fluid ability tests measure primarily metacomponential functioning, performance-componential functioning (and particularly, the performance components of induction), and the application of these various components to relatively novel situations. The tasks may or may not be novel to particular individuals, depending on their past history and experience with such tasks. Crystallized ability tests measure primarily the products, rather than the processes of knowledge acquisition. They are thus indirect measures of componential functioning. I believe it is because these tests emphasize the measurement of products rather than processes that they are relatively more immune to age-related decline than are fluid ability tests (Horn, 1968).

Guilford's (1967) theory of intelligence has always been something of an anomaly. The factor-analytic evidence that has been offered in favor of it is of questionable validity (Horn & Knapp, 1973), and hence it is easy to dismiss the theory. But Guilford's theory makes certain important contributions that are

unusual in factor-analytic theories. First, it explicitly builds the notion of process into the factors of intelligence, something no previous theory had done. Although the various postulated processes may not all be independent psychometric factors, they are at least recognized as distinctly contributing to the factor model. In the triarchic theory, of course, processes are also accorded a major role in understanding the nature of intelligence. Second, Guilford was among the earliest theorists to recognize the importance of everyday behavioral competence in intelligence. His theory was not limited in its scope merely to the academic side of intelligence. This emphasis appears in the contextual sub-theory of the triarchic theory. Finally, the theory explicitly includes, within the structure-of-intellect cube, abilities relevant to dealing with novelty, in particular, the divergent thinking abilities. Although I question the validity of some of the tests Guilford has used to measure this construct, I applaud his recognition of the importance of dealing with novelty to intelligence. There are other positive features of Guilford's theory that could be mentioned, but I would hope the point is clear: Whether or not the theory is correct in its details (which I doubt), it is one of the more progressive psychometric theories of intelligence to have been proposed.

Information-processing theories of intelligence differ primarily in the level of information processing that they emphasize. As noted earlier, theories such as Jensen's (1982) emphasize lower levels of processing, whereas theories such as Sternberg's (1980) or Snow's (1979) emphasize higher levels of processing. In the triarchic theory, intelligence is understood in terms of the interaction of all of these levels. Hence, a complete theory of intelligence would have to account for individual differences in choice reaction time as well as individual differences in complex problem solving, as in complex analogies. This is not to say that all levels of processing contribute equally to individual differences in all societies, at all ages, or in all possible groups. I doubt this to be the case. But the question here is of degree of componential contribution rather than of kind of componential contribution, and also of the degree to which the execution of the various component processes is automatized. The various information-processing theories serve a useful function in highlighting the various levels of analysis that are possible in understanding intelligent processing of information.

Piaget's theory of intelligence (e.g., Piaget, 1972) has so many aspects that it would not be possible to deal with all of them here. The richness of this theory is probably unmatched in any other theory of intelligence. I will dwell here on only two aspects of the theory, the mechanisms of equilibration and the stages of development.

In Piaget's theory, adaptation to the environment occurs through two mental mechanisms: assimilation and accommodation. In assimilation, one fits a new stimulus into one's existing mental structures. In accommodation, one alters one's mental structures in order to understand the new stimulus. The mecha-

nism of equilibration (i.e., the balance between assimilation and accommodation) thus highlights the role of dealing with novelty in ecologically valid situations. The triarchic theory does not have two processes that directly correspond to assimilation and accommodation. Rather, it has three components of knowledge acquisition—selective encoding, selective combination, and selective comparison—that can be applied to tasks and situations at differing levels of novelty and cognitive preparedness. Whether or not a given process results in a new cognitive structure is viewed as a representational issue rather than one of process: In the triarchic theory, application of a process of knowledge acquisition may or may not change one's mental representation or structures without changing the process involved. But this theory, like Piaget's, deals with the aspect of intelligence involving learning new information in tasks and situations that are relatively novel, and that are ecologically valid.

The triarchic theory does not postulate stages of intellectual development, although it does postulate a series of mechanisms by which cognitive development takes place (see Sternberg, 1985a). I doubt that anything like discrete stages exist in cognitive development, and the bulk of recent evidence seems consistent with this doubt (see, e.g., Gelman & Baillargeon, 1983). In the triarchic theory, it is possible for sets of components to become available or increasingly accessible at about the same time, and such increases in availability and accessibility might render an appearance, at least in a rough sense, of a stagelike progression of cognitive development.

Some theorists of intelligence have emphasized the role of environmental context both in determining what intelligence is and in shaping one's level and kinds of intelligence (e.g., Berry, 1972; Laboratory of Comparative Human Cognition, 1982). These theories are partially consistent with the contextual subtheory of the triarchic theory, which also emphasizes how context shapes intelligence and intelligence shapes context. If there is a difference, it is in the triarchic position, shared by Irvine (1979), that there are constancies in intelligence across cultures that can help protect us from a position of total relativism.

CONCLUSIONS AND IMPLICATIONS

The triarchic theory consists of three interrelated subtheories that attempt to account for the bases and manifestations of intelligent thought. The componential subtheory relates intelligence to the internal world of the individual. The experiential subtheory relates intelligence to the experience of the individual with tasks and situations. The contextual subtheory relates intelligence to the external world of the individual.

The elements of the three subtheories are interrelated: The components of intelligence are manifested at different levels of experience with tasks and in

situations of varying degrees of contextual relevance to a person's life. The components of intelligence are posited to be universal to intelligence: Thus, the components that contribute to intelligent performance in one culture do so in all other cultures as well. Moreover, the importance of dealing with novelty and automatization of information processing to intelligence are posited to be universal. But the manifestations of these components in experience are posited to be relative to cultural contexts. What constitutes adaptive thought or behavior in one culture is not necessarily adaptive in another culture. Moreover, thoughts and actions that would shape behavior in appropriate ways in one context might not shape them in appropriate ways in another context. Finally, the environment one selects will depend largely upon the environments available to one, and the fit of one's cognitive abilities, motivations, values, and affects to the available alternatives.

The triarchic theory has certain implications both for the assessment and the training of abilities. With regard to assessment, a full assessment battery would necessarily tap all of the abilities specified by the triarchic theory, something no existing test even comes close to. Although I am pursuing the development of a triarchic test of intelligence, even a test explicitly designed to measure intelligence according to the triarchic theory will be only an approximation to an ideal test, if only because the relativity of the contextual subtheory renders any one test adequate only for a limited population. Similarly, I have developed a training program for understanding and improving intellectual skills based on the theory (Sternberg, 1986). But the training program could not possibly develop all of the skills posited by the theory, especially because contextual skills can be so variable across environments.

In conclusion, the triarchic theory offers a relatively complete account of intelligent thought that draws upon and partially subsumes many existing theories. This new theory, like all other theories, is only an approximation, one that will serve a constructive purpose if it, too, is eventually subsumed by a more complete and accurate theory of human intelligence. But it does not deal with styles in the way intelligence is actually used. For such an issue, one needs a separate theory, one of intellectual styles.

INTELLECTUAL STYLES THREE PROFILES OF STYLES

Throughout my four years in college, my two roommates and I remained together. The roommates—Alex, Bob, and Cyril (only one of these names is unchanged)—seemed remarkably similar intellectually when they entered college. All had high Scholastic Aptitude Test scores, excellent grades in high school, and similar intellectual strengths and weaknesses. For example, all

three were more verbal than quantitative, reasoned well, but were rather weak spatially. Thus, in terms of standard theories of intelligence, the three roommates had similar intellectual abilities. Moreover, today, all three roommates are successful in their jobs, and have achieved some national recognition for their work, showing that the three roommates were similar in motivational levels as well.

If one looks beyond the intellectual similarities of the three roommates, one cannot help but notice some salient differences that have profoundly affected their lives. Consider some of the differences between Alex, Bob, and Cyril.

Alex, a lawyer, is admittedly fairly conventional, rule-bound, and comfortable with details and structure. He does well what others tell him to do, as a lawyer must, and has commented to me that his idea of perfection would be a technically flawless legal document or contract whereby those who sign on the dotted line are bound to the terms of the contract without loopholes. In a nutshell, Alex is a follower of systems, and follows them extremely well, as shown by the facts that he was formerly a Rhodes Scholar and is today a partner in a major national law firm. Alex can figure out a system and work excellently within it.

Bob, a university professor, is quite different stylistically from Alex. He is fairly unconventional, and unlike Alex, dislikes following or even dealing with other people's rules. Moreover, he has relatively few rules of his own. Although he has some basic principles that he views as invariants, he tends not to take rules very seriously, viewing them as conveniences that are meant to be changed or even broken as the situation requires. Bob dislikes details, and generally is comfortable working within a structure only if it is his own. He does certain things well, but usually only if they are the things he wants to do, rather than what someone else wants him to do. His idea of intellectual perfection would be the generation of a great idea and a compelling demonstration that the idea is correct, or at least, useful. In brief, Bob is a creator of systems, and has designed some fairly well-known psychological theories that reflect his interest in system creation.

Cyril, a psychotherapist, is like Bob but not Alex in being fairly unconventional. Like Bob, he dislikes others' rules, but unlike Bob, he has a number of his own. He tends to be indifferent to details. He likes working within certain structures, which need not be his own, but the structures have to be ones he has adjudged to be correct and suitable. Cyril does well what he wants to do. His idea of perfection would be a difficult but correct psychological diagnosis, followed by an optimal psychotherapeutic intervention. In sum, Cyril is a judge of systems. His interest, perhaps passion, for judging was shown early in his career, when, as a college student, he constructed a test (which we called the "Cyril Test") to give to others, and especially to dates, to judge the suitability

of their values and standards. Cyril was also editor of the college course critique, a role in which he took responsibility for the judgment and evaluation of all undergraduate courses at the university.

Although Alex, Bob, and Cyril are all intellectually able and similarly competent, even these brief sketches serve to illustrate that they use their intelligence in different ways. Alex is a follower or executor, Bob, a creator or legislator, and Cyril, a judge of systems. They differ in terms of their intellectual style, or ways in which they direct their intelligence. A style, then, is not a level of intelligence, but a way of using it—a propensity. When one is talking about styles rather than levels, one cannot talk simply about better or worse. Rather, one must speak in terms of better or worse “for what?”

THE MODEL OF INTELLECTUAL STYLES AS MENTAL SELF-GOVERNMENT

I am proposing here a model of intellectual styles as mental self-government (Sternberg, 1988). The basic idea is that governmental structures may be external, societal manifestations of basic psychological processes that are internal and individual (see also Bronfenbrenner, 1977). Seeds of this notion can be found in the writings of political theorists, such as Plato, Hobbes, Locke, and Rousseau, whose political theories were based on psychological theories of what people are like. The difference here, perhaps, is that rather than attempting to understand governments in terms of the psychology of human beings, we try to understand the psychology of human beings in terms of governments. From this point of view, government in society is a large-scale, externalized mirror of the mind. People are systems, just like societies (Ford, 1986), and they need to govern themselves just as do societies. Mental incompetence results from a breakdown of self-regulating functions, and high levels of mental competence derive in part from superior self-regulation.

The view of intellectual styles as mental self-government focuses more on styles than on levels of intelligence. In standard theories of intelligence, including recent ones (Gardner, 1983; Sternberg, 1985a), the emphasis is on levels of intelligence, of one or more kinds. Measuring intelligence thus entails assessing how much of each ability the individual has. In contrast, the governmental model leads to assessment not of how much intelligence the individual has, but rather of how that intelligence is directed, or exploited. Two individuals of equal intelligence, by any of the existing theories of intelligence, might nevertheless be viewed by the theory as intellectually quite different because of the ways in which they organize and direct that intelligence. In the next part of the chapter, the implications of the mental self-government model as a basis for understanding intellectual styles are explored in some detail.

Governments have many aspects, such as function, form, level, scope, and leaning. Three major functions of government are the legislative, executive, and judicial. Four major forms of government are the monarchic, hierarchic, oligarchic, and anarchic. Two basic levels of government are the global and the local. Two domains in the scope of government are the internal (domestic affairs) and the external (foreign affairs), and two leanings are conservative and progressive. In this part of the chapter, the implications of each of these aspects for understanding intellectual styles are explored. People do not have—exclusively—one style or another, but rather, show preferences for styles that result in different preferred ways of using their intelligence.

The Functions of Government

Governments may be viewed as having three primary functions: legislative, executive, and judicial.

The legislative style characterizes individuals who enjoy creating, formulating, and planning for problem solution. Such individuals, like Bob, the university professor described earlier, like to create their own rules, enjoy doing things their own way, prefer problems that are not prestructured or prefabricated, and like to build structure as well as content in deciding how to approach a problem. People with legislative tendencies prefer creative and constructive planning-based activities, such as writing papers, designing projects, and creating new business or educational systems. They tend to enter occupations that enable them to utilize their legislative style, such as creative writer, scientist, artist, sculptor, investment banker, policy maker, and architect.

Individuals with an executive style are implementers. Like Alex, the lawyer described earlier, they like to follow rules and work within existing systems, preferring prestructured or prefabricated problems which allow them to fill in content within existing structures. They prefer predefined activities such as solving algebra-word problems or engineering problems, giving talks or lessons based on others' ideas, and enforcing rules. Executive types gravitate toward occupations such as lawyer, policeman, builder (of others' designs), surgeon, soldier, proselytizer (of others' systems), and manager (lower echelon).

The judicial style, as shown by the psychotherapist, Cyril, described earlier, involves judgmental activities. Judicial types like to analyze and criticize, preferring problems in which they evaluate the structure and content of existing things and ideas. They prefer activities that exercise the judicial function, such as writing critiques, giving opinions, judging people and their work, and evaluating programs. People with a primarily judicial style tend to gravitate toward occupations such as judge, critic, program evaluator, admissions officer, grant or contract monitor, systems analyst, and consultant.

People do not have one or another style exclusively—rather, they tend to specialize, some people more than others. For example, one individual might be strongly legislative and only weakly executive and judicial, whereas another individual might be approximately equally balanced among the three functions. Thus, people differ not only in their direction of specialization, but in the degree to which they specialize. People will gravitate toward problems whose solutions require the people's preferred styles of functioning. They may also use certain styles in the service of other styles. A primarily legislative type, for example, may use judicial functions primarily to further legislative ends.

We need to distinguish the proclivity toward a style from people's abilities to implement that style. It seems likely that most people will prefer styles that capitalize upon their strengths. But there is no logical or psychological reason why preferences and abilities will always correspond. Some people may prefer styles that are not as well suited to their abilities as are others. In measuring styles, it is important to measure both predilections toward styles and abilities to implement them, in order to determine how well an individual's predilections and abilities match.

An important implication of these differences is that although style is generally independent of level of intelligence, it probably is not independent of level of *perceived* intelligence within a particular domain. The same individual who might be thought to be a brilliant science student because he is a legislative type might be thought to be somewhat duller in business courses that more emphasize executive skills.

Forms of Mental Self-Government

Governments come in different forms. Four of those forms are the monarchic, the hierarchic, the oligarchic, and the anarchic. Logically, any form may be paired with any function, although psychologically, certain pairings are likely to be more common than others. People prefer to organize their experience in ways that correspond to the various forms of government.

People who exhibit a predominantly monarchic style tend to be motivated by a single goal or need at a time. Single-minded and driven, they often believe that the ends justify the means and attempt to solve problems, full-speed ahead—damn the obstacles. They may tend to oversimplify problems, often being more decisive than the situation warrants. In a limited sense they may be systematic, however, they may neglect variables not obviously pertinent to their goal.

Individuals preferring a hierarchic style tend to be motivated by a hierarchy of goals, with the recognition that not all goals can be fulfilled equally well and that some goals are more important than others. They take a balanced approach to problems, believing that ends do not justify means, and viewing

competing goals as acceptable (although they may have trouble if the priorities come too close to allow for formation of a hierarchy). Hierarchic types seek complexity and tend to be self-aware, tolerant, and relatively flexible. They have good senses of priorities, are usually decisive, unless priority setting becomes a substitute for decision or action, and are systematic in problem solving and decision making.

Individuals preferring the oligarchic style tend to be motivated by multiple, often competing goals of equal perceived importance. Plagued by multiple, possibly competing approaches to problems, they are often driven by goal conflict and tension arising out of their belief that satisfying the constraints is as important as solving the problem itself. They usually believe that ends do not justify means, and find that competing goals and needs tend to interfere with task completion, because each goal and need is seen as of roughly equal importance. Oligarchic types seek complexity (sometimes to the frustration point), and are self-aware, tolerant, and very flexible. They tend to have trouble setting priorities because everything seems equally important, and thus they are rather indecisive and multiply systematic, with the multiple systems competing with each other because of the need to satisfy multiple equally important goals.

Anarchic stylists tend to be motivated by a potpourri of needs and goals that are often difficult for themselves, as well as others, to sort out. They take a random approach to problems, driven by what seems to be a muddle of inexplicable forces. They may act as though ends justify means. They may believe that anything goes, and have trouble setting priorities because they have no firm set of rules upon which to base them. They tend to be extreme, being either too decisive or too indecisive, and are thoroughly asystematic.

Some general issues arise with regard to formal style of mental self-government. Monarchists may be too single-minded for the likes of most teachers and even social acquaintances. But in later life, their single-minded zeal may render them among the most successful of entrepreneurs or goal attainers. Often, their memories of school will not be fond, because they will believe that their talents went unrecognized. Monarchists can also be difficult to live with because of their single-mindedness.

Hierarchical types can probably solve the widest variety of problems in school life and beyond, because most problems are probably best conceived of hierarchically. They will generally achieve a good balance between thought and action, but they must remember that the existence of priorities does not guarantee that those priorities are right. When there is a serious bottom line, or pressing goal, hierarchists may get lost or sidetracked in their own hierarchies, whereas the monarchist may blitz through and attain the goal.

Oligarchists will often frustrate themselves and others, in school and in careers, because of their indecision and hesitation. Because they tend to assign equal weights to competing means and goals, they may appear to be "lost in

thought” and unable to act. They can act, but they may need others to set their priorities for them.

Anarchists are at risk of becoming educational as well as social misfits, and their talents may actually lead them into anti- rather than prosocial paths. Properly nurtured, they may have the potential for truly creative contributions to the world, if their anarchic style is combined with the necessary intellectual talents for creative performance. But proper nurturance may be quite a challenge because of the anarchists’ unwillingness to work within existing systems in order eventually to go beyond these systems. Rather than working within existing systems, anarchists may end up attempting to destroy them.

Levels of Mental Self-Government

Globalists prefer to deal with relatively large and abstract issues. They tend to ignore or dislike detail, choosing instead to conceptualize and work in the world of ideas. Their weaknesses are that they may be diffuse thinkers, who can get lost on “Cloud 9,” and that they may see the forest but not always the trees within it.

In contrast, localists like concrete problems requiring detailed work, and are often pragmatically oriented and down to earth. Their weakness, however, is that they may not see the forest for the trees.

In terms of the three individuals described earlier—Alex, Bob, and Cyril—Bob and Cyril tend to be globalists whereas Alex tends to be a localist. The local style is not, however, inextricably linked to the executive style Alex has shown. Some executive types may prefer only to work at a broader level, accomplishing the main tasks in a project while relegating the more local details to others. Similarly, a legislative or judicial type could be more local than either Bob or Cyril.

Although most people prefer to work either at a more global or a more local level, a key to successful problem solving in many situations is being able to traverse between levels. If a person is weak within a given level, it is often helpful to pair up with someone whose strengths are complementary. In particular, although we often value most people who are most like ourselves, we actually benefit most from people who are moderately unlike ourselves with respect to preferred level of processing. Too much overlap leads to some levels of functioning simply being ignored. Two globalists, for example, may do well in forming ideas, but will need someone to take care of the details of implementing them. Two localists may help each other in implementation, but need someone to set down the global issues that need to be dealt with. Too little overlap, however, can lead to a breakdown in communication. People who do not overlap at all in levels may not be able to understand each other well.

Scope of Mental Self-Government

Governments may be internally oriented and isolationist, or externally oriented and collaborative with the other nations. Similarly, mental self-governments need to work both alone and with others.

Internalists tend to be introverted, task-oriented, aloof, socially less sensitive, and interpersonally less aware than externalists. They also like to work alone. Essentially, their preference is to apply their intelligence to things or ideas in isolation from other people.

Externalists tend to be extroverted, people-oriented, outgoing, socially more sensitive, and interpersonally more aware than internalists. They like to work with others, and seek problems that either involve working with other people or are about others.

Among the three individuals described earlier, Alex and Bob tend more toward the internal scope of mental self-government whereas Cyril tends more toward the external. These proclivities fit with their jobs. Alex works primarily in corporate law, dealing with legal principles and documents and less with people; Bob works primarily with ideas and instantiating them through experiments. Cyril, as a psychotherapist, is constantly working with people. It should be realized that there is some degree of situation specificity involved. Bob, for example, works actively with students and frequently gives lectures on his work. At the same time, he tends to shun parties, and generally prefers to deal with people socially most when there is at least some degree of task orientation. Moreover, he recognizes the importance of dealing with people on his job, and makes sure that whatever his preferred tendencies, the job gets done when interactions with people are required.

Some people prefer to be internalists whereas others prefer to be externalists. Again, most people are not strictly one or the other, but alternate between them as a function of task, situation, and the people involved. But it is important to realize in education and job placement that a bright individual who is forced to work in a mode that does not suit him or her may perform below his or her capabilities.

Leanings of Mental Self-Government

Governments can have various leanings. For present purposes, two major "regions" of leanings will be distinguished, conservative and progressive.

Individuals with a predominantly conservative style like to adhere to existing rules and procedures, minimize change, and avoid ambiguous situations where possible, preferring familiarity in life and work.

Individuals with a progressive style like to go beyond existing rules and procedures, maximize change, seek or at least accept ambiguous situations, and prefer some degree of unfamiliarity in life and work.

Although individuals may, on the average, tend toward a more conservative or progressive leaning in their mental self-government, there is clearly some degree of domain specificity involved. For example, an individual who is conservative politically will not necessarily be conservative in his or her personal life, and similarly for a progressive. Thus, in evaluating styles, and especially leanings, tendencies within particular domains must be taken into account. Moreover, leanings may well change over time as people feel more or less secure in their environments. Thus, an individual who is new to an environment may tend to adapt conservatively, whereas an individual who has been in that environment longer may feel more free progressively to attempt to shape the environment. This aspect of style may be among the most mercurial of the various aspects.

DEVELOPMENT OF INTELLECTUAL STYLES

Where do these various modes of intellectual functioning come from? It is possible that at least some portion of stylistic preference is inherited, but I doubt that it is a large part. Rather, styles seem to be partly socialized constructs, just as is intelligence (Sternberg & Suben, 1986). From early on, we perceive certain modes of interaction to be rewarded more than others, and we probably gravitate toward these modes, while being constrained by our built-in predispositions as to how much and how well we are able to adopt these rewarded styles.

Consider some of the variables that are likely to affect the development of intellectual styles.

A first variable is culture. Different cultures tend to reward different styles. For example, the North American emphasis on innovation ("making the better mouse trap") may lead to relatively greater reward for the legislative and progressive styles, at least among adults. National heroes in the United States, such as Edison as inventor, Einstein as scientist, Jefferson as political theorist, Steve Jobs as entrepreneur, and Hemingway as author, tend often to be heroes by virtue of their legislative contribution. Other societies that tend to value conformity and the following of tradition may be more likely to reward executive and conservative styles. A society that emphasizes conformity and tradition to too great a degree may stagnate because of the styles induced into its members.

A second variable is gender. Traditionally, a legislative style has been more acceptable in males than in females. Men were supposed to set the rules, and

women to follow them. Although this tradition is changing, the behavior of many men and women does not fully reflect the new values.

A third variable is age. Legislation is generally encouraged in the preschool young, who are encouraged to develop their creative powers in the relatively unstructured and open environment of the preschool and some homes. Once the children start school, the period of legislative encouragement rapidly draws to a close. Children are now expected to be socialized into the largely conforming values of the school. The teacher now decides what the student should do, and the student does it, for the most part. Students who don't follow directions and the regimentation of the school are viewed as undersocialized and even as misfits. In adulthood, some jobs again encourage legislation, even though training for such jobs may not. For example, high school physics or history are usually largely executive, with students answering questions or solving problems that the teacher poses. But the physicist and historian are expected to be more legislative. Ironically, they may have forgotten how. We sometimes say that children lose their creativity in school. What they may really lose is the intellectual style that generates creative performance.

A fourth variable is parenting style. What the parent encourages and rewards is likely to be reflected in the style of the child. Does the parent encourage or discourage legislation, or judgment, on the part of the child? The parent him or herself exhibits a certain style, which the child is likely to emulate, and modeling is an important source of behavioral development. A monarchic parent, for example, is likely to reward a child who shows the same single-mindedness, whereas an anarchic parent would likely abhor a child beginning to show a monarchic style, and try to suppress it as unacceptable. Parents who mediate for the child in ways that point to larger rather than smaller issues underlying actions are more likely to encourage a global style, whereas parents who do not themselves generalize are more likely to encourage a more local style. And mediation, too, is a powerful source of development (Feuerstein, 1980).

A last variable is kind of schooling and, ultimately, of occupation. Different schools and, especially, occupations reward different styles. An entrepreneur is likely to be rewarded for different styles from those for which an assembly-line worker is rewarded. As individuals respond to the reward system of their chosen life pursuit, various aspects of style are more likely to be either encouraged or suppressed.

Obviously, these variables are only a sampling rather than a complete listing of those variables that are likely to influence style. Moreover, any discussion such as this one inevitably simplifies the complexities of development, if only because of the complex interactions that occur among variables. Moreover, styles interact with abilities. Occasionally one runs into legislative types who are uncreative, creative people who eschew legislation, hierarchists

who set up misguided hierarchies, and so on. But for the most part, the interactions will be more synchronous in well-adjusted people. According to the triarchic theory of human intelligence (Sternberg, 1986), contextually intelligent people are ones who capitalize on their strengths and who either remediate or compensate for their weaknesses. A major part of capitalization and compensation would seem to be in finding harmony between one's abilities and one's preferred styles. People who cannot find such harmony are likely to be frustrated by the mismatch between how they want to perform and how they are able to perform.

If styles are indeed socialized, even in part, then they are almost certainly modifiable to at least some degree. Such modification may not be easy. We know little about how to modify intelligence, and we know even less about how to modify intellectual styles. Presumably, when we learn the mechanisms that might underlie such attempts at modification, we will pursue a path similar to that which some educators and psychologists are using in teaching intelligence (e.g., Sternberg, 1986).

We need to teach students to make the best of their intellectual styles. Some remediation of weaknesses is probably possible. But to the extent that it is not, mechanisms of compensation can usually be worked out that help narrow the gap between weak and strong areas of performance. For example, groups of children can be formed that pair children with different preferred styles. Ultimately, we can hope that a theory of intellectual styles will serve not only as a basis for a test of such styles, but also as a basis for training that maximizes people's flexibility in dealing with their environment, society, and themselves. The main way of teaching may well end up being role modeling of teachers by learners.

Schools most reward executive types—children who work within existing rule systems and seek the rewards that the schools value. To some extent, the schools create executive types out of people who might have been otherwise. But whether the rewards will continue indefinitely for the executive types depends in part upon career path, which is why school grades are poor predictors of job success. One's ability to get high grades in science courses involving problem solving, for example, probably will not be highly predictive of one's success as a scientist, an occupation in which many of the rewards are for coming up with the ideas for the problems in the first place. Judicial types may be rewarded somewhat more in secondary and especially tertiary schooling, where at least some judgmental activity is required, as in paper writing. Legislative types, if they are rewarded at all, may not be rewarded until graduate school, where there is a need to come up with one's own ideas in dissertation and other research. But some professors—those who want students who are clones or disciples—may not reward legislative types even in graduate

school, preferring executive types who will carry out their research for them in an effective, diligent, and nonthreatening way.

The fit between student and teacher, as between principal and teacher, can be critical to the success of the teacher–student system, or of the principal–teacher system. A legislative student and an executive teacher, for example, may not get on well at all. A legislative student may not even get along with a legislative teacher if that teacher happens to be one who is intolerant of other people’s legislations. During the course of my career, I have found that although I can work with a variety of students, I probably work best with students whom I now, in retrospect, would classify as legislative. I can work reasonably well with executive types also. I am probably weakest with judicial students, who to me seem more eager to criticize than to do research. The general point is that educators need to take into account their own styles in order to understand how they influence their perceptions of and interactions with others. Clearly, certain children benefit from certain styles. A gifted executive-type student might benefit more from acceleration, where the same material is presented at a more rapid pace. A gifted legislative-type student might benefit more from enrichment, where the opportunity to do creative projects would be consistent with the student’s preferred style of working.

It is necessary that schools take into account not only fit between teacher and student (or principal and teacher) style, but also the fit between the way a subject is taught and the way a student thinks. A given course often can be taught in a way that is advantageous (or disadvantageous) to a particular style. Consider, for example, an introductory or low-level psychology course. This course might stress learning and using existing facts, principles, and procedures (an executive style of teaching), or it might stress designing a research project (a legislative style of teaching), or it might stress writing papers evaluating theories, experiments, and the like (a judicial style of teaching). Little wonder I received a grade of “C” in my introductory psychology course, taught in the executive style! And, in retrospect, little wonder that in my own psychology courses, I have almost always made the final grade heavily dependent on the design of a research project. My style of teaching was reflecting my own style of thinking, as it does for others. The general principle of style of teaching reflecting the teacher’s preference is not limited to psychology or even science. Writing, for example, might be taught in a way that emphasizes critical (judicial) papers, creative (legislative) papers, or expository (executive) papers.

Sometimes, there is a natural shift in the nature of subject matter over successive levels of advancement, just as there is in jobs. In mathematics and basic science, for example, lower levels are clearly more executive, requiring solution of prestructured problems. Higher levels are clearly more legislative, requiring formulation of new ideas for proofs, theories, and experiments.

Unfortunately, some of the students screened out in the earlier phases of education might have succeeded quite well in the later ones, whereas some students who readily pass the initial stages might be ill-suited to later demands.

Perhaps the most important point to be made is that we tend to confuse level with style of intelligence. For example, most current intelligence and achievement tests reward the executive style by far the most—they require solution of prestructured problems. One cannot create one's own problems, or judge the quality of the problems on the test (at least not at the time of test!). Judicial types get some credit for analytical items, but legislative types hardly benefit at all from existing tests, and may actually be harmed by them. Clearly, style will affect perceived competence, but as noted earlier, style is independent of intelligence, in general, although not within particular domains. Style ought to count as much as ability and motivation in recommending job placements, although probably not in making tracking decisions that deal with issues of ability rather than style.

The styles of intellect proposed here are not, of course, the only ones ever to have been proposed. Theories of intellectual styles abound, and although it is not possible to review them exhaustively here, I will cite some pertinent examples.

Myers (1980; see also Myers & McCaulley, 1985) has proposed a series of psychological types based on Jung's (1923) theory of types. According to Myers, there are 16 types, resulting from all possible combinations of two ways of perceiving—sensing versus intuition; two ways of judging—thinking versus feeling; two ways of dealing with self and others—introversion versus extraversion; and two ways of dealing with the outer world—judgment versus perception.

Gregorc (1985) has proposed four main types or styles, based on all possible combinations of just two dimensions—concrete versus abstract and sequential versus random.

Concrete-sequential refers to a preference for the ordered, the practical, and the stable. Individuals who are dominant concrete-sequentials have a tendency to focus their attention on material reality and physical objects and to validate ideas via the senses. These individuals also have a tendency to conform. Abstract-sequential refers to a preference for mentally stimulating environments. Individuals who are dominant abstract-sequentials have a tendency to focus their attention on the world of the intellect. They are characterized by a preference for logical and synthetic thinking and for validating information via personal formulae. Abstract-random refers to a preference for emotional and physical freedom. Individuals who are dominant abstract-randoms have a tendency to focus their attention on the world of feeling and emotion. They are also characterized by a tendency to validate ideas via inner guidance. Finally, concrete-random refers to a preference for a stimulus-rich environment that is free from restriction. Individuals who are dominant concrete-randoms tend to

engage with the concrete world of activity and the abstract world of intuition. They tend to prefer intuitive and instinctive thinking and to rely on personal proof for validating ideas, rarely accepting outside authority.

Taking a more educationally oriented slant, Renzulli and Smith (1978) have suggested that individuals have various learning styles, with each style corresponding to a method of teaching: projects, drill and recitation, peer teaching, discussion, teaching games, independent study, programmed instruction, lecture, and simulation. Holland (1973) has taken a more job-related orientation and proposed six styles that are used as a basis for understanding job interests as revealed by the Strong-Campbell Interest Inventory (1985). Holland's typology includes six "types" of personality: realistic, investigative, artistic, social, enterprising, and conventional.

Intellectual styles represent an important link between intelligence and personality, because they probably represent a way in which personality is manifested in intelligent thought and action. Attempts to understand academic or job performance solely in terms of intelligence or personality probably have not succeeded as well as we had hoped because they neglect the issue of intellectual style—the effect of intelligence and personality on each other. Thus, styles may represent an important "missing link" between intelligence, personality, and real-world performance. The basic idea is that just as we have theories of how modifiable individuals are, we need a theory of the modifiability of contexts. People talk a lot about context, but there is little in the way of theories of context, such as that to be presented here. In times past, we used to treat variation across individuals (i.e., individual differences) as "error." Some people still do. Now we treat variation across context as error. For example, if a program works in one school but not in another, we write that off as the "error term" in our analysis. But those who have worked with schools know that interventions in schools work in some school contexts but not in others. Why does this happen? Is there any way of knowing in advance how likely the context is to be modifiable, independent of the specific intervention proposed for it?

The theory of contextual modifiability requires that one ask three questions about a context in order to assess its modifiability. The three questions are:

1. How much desire is there for actual change in this context as a whole?
2. How much desire is there for the *appearance* of change in the context?
3. What is the self-esteem, or opinion of itself, of the context as a whole?

If, for convenience, we respond to each of these questions with a value that is either "low" or "high," then we end up with 2^3 , or eight different kinds of contexts with respect to modifiability. The argument here is that the eight kinds of contexts differ rather dramatically in how modifiable they are. Of course, a context need not be a pure case: It may be a mixture. Moreover, one would

certainly not want to claim that these are the only possible types of contexts with respect to modifiability. Yet, they do appear to encompass some of the major types. Indeed, they could be applied in theory to any context, whether at the level of the organization, organizational unit, or even of the individual.

Modifiability can be of two basic kinds: surface-structural or deep-structural. (Of course, there is an underlying continuum here which I have dichotomized for convenience.) Surface-structural modifiability refers to the extent to which a modification seeks merely to build on what is already there. Surface-structural modifications make changes, but within the context as it exists. Deep-structural modifiability refers to the extent to which more profound changes can be achieved. A deep-structural intervention requires building new structures and making at least some fundamental changes in the nature of the context. Modifiability depends primarily on desire for actual change, and secondarily on desire for appearance of change and self-esteem. Theory is presented at a general level, which could apply to a school district or even a teacher as well as to a school. For convenience, the unit of analysis here will be the school. Contexts may sometimes be at least partially domain-specific. In other words, one department will fall into one category, another department (or other unit) into another category. Moreover, modifiability may also depend on whether what an individual has to offer matches the context's priorities for change. The context may be modifiable with respect to some areas but not others.

Each of the eight types of context is depicted in terms of a different kind of mineral. Table 1 summarizes the theory of contextual modifiability.

ALTERNATIVE CONTEXTS AND THEIR MODIFIABILITY

The Rusted-Iron School

The Rusted-Iron School is low in desire for actual change, desire for appearance of change, and self-esteem. The mood of the school is despondence. Its self-belief is that "We're lost. We're hopeless." Typical self-statements of this school are: "We were once OK, but now we're long gone"; "We're beaten down with bureaucracy from one side, and discipline problems from the other"; and "We just try to get through the day."

Various signs indicate a Rusted-Iron School. Among these signs are an entrenched bureaucracy, apathy, a decayed physical plant, staff burnout, lack of follow-through on agreements, indifference to served population, and lack of resources. The prognosis for change in a school showing these signs is poor: The likelihood of both surface-structural and deep-structural change is low.

Table 1. A Theory of Contextual Modifiability

<i>DESIRE FOR ACTUAL CHANGE</i>	<i>DESIRE FOR APPEAR- ANCE OF CHANGE</i>	<i>SELF ESTEEM</i>	<i>MODIFIABILITY</i>		<i>DESCRIPTION</i>
			<i>Surface Deep</i>		
L	L	L	L	L	RUSTED IRON
L	L	H	L	L	GRANITE
L	H	L	ML	L	AMBER (with Internal Insects)
L	H	H	ML	L	OPAL
H	L	L	ML	L	CUBIC ZIRCONIUM
H	L	H	MH	ML	SLIGHTLY IMPERFECT (SI) DIAMOND
H	H	L	MH	ML	LEAD
H	H	H	H	H	DIAMOND IN THE ROUGH

L = Low; M = Medium; H = High

The Granite School

The Granite School is low in desire for actual change, low in desire for appearance of change, but high in self-esteem. Its mood is one of smugness. Its self-belief is that “We’re sure and solid like the Rock of Gibraltar. Change would only chip away at us.” Typical self-statements are: “We may not look great, but we’re solid and durable”; “This is a school that works”; and “We teach students the basics, and we keep them in line.”

Some signs of a Granite School are a traditional curriculum, emphasis on discipline, pride in always having done things the way they are now being done, old materials, and grimness in the attitude of personnel. The prognosis for this school, like the prognosis for the Rusted-Iron School, is poor: Both surface-structural and deep-structural modifiability are low.

The Amber School (with Internal Insects)

The Amber School is low in desire for actual change, high in desire for the appearance of change, and low in self-esteem. Its mood is one of frustration. Its self-belief is that “We’re internally flawed. To change would destroy our very core and us with it.” Thus, the Amber School believes that it has internal flaws: If they were to be removed, it would result in the destruction of the school, just as removing internal insects from amber would destroy the amber. Typical self-statements are: “We know we’ve got problems, but you just can’t beat the system here”; “Moving the school ahead is like moving a graveyard

anywhere”; and “The core is rotten: The administration hasn’t budged in years.”

Some signs of an Amber School are an inured and often aging administration, obvious structural flaws in the instructional program, hyperstability in the face of dissension, and inaccessibility of the power structure. The prognosis for surface change is medium low, and the prognosis for deep-structural change is low.

The Opal School

The Opal School is low in desire for actual change, but high in desire for appearance of change and high in self-esteem. Its mood is one of self-righteousness. Its typical self-belief is that “When you’re the best, you’ve got to put your efforts into staying that way.” Typical self-statements are: “Anywhere you look, creative, new things are happening”; “We’re the best in the state”; and “Look at this gym (or lab or auditorium or whatever).”

This school is like an opal in that if you look at it from different perspectives, it looks different and like it’s changing, when in fact, it’s always the same. Only the appearances are different. And the power structure of the school believes that changing the school, like changing an opal, is likely only to make it worse. It fares best when left alone.

Signs of an Opal School are typically affluence, a shiny physical plant, many (often unused or ill-used) resources, slickness of administrators, clear emphasis on appearances, high salaries, and surprising lack of mission. The prognosis for surface-structural change is moderately low, the prognosis for deep-structural change, low.

The Cubic Zirconium School

The Cubic Zirconium School is high in desire for actual change, but low in both desire for the appearance of change and in self-esteem. The mood of the school is fraudulence: As is the case with a cubic zirconium, no one wants viewers to know that it is fraudulent, so viewers are kept at a distance. Its self belief is that “We’re a fraud; We can’t let outsiders get too close, lest they find out.” Typical self-statements are: “We can’t have outsiders disrupt our educational program,” “You can see yourself that things are fine here. Thanks for stopping by,” and “We don’t do research here; we teach children.”

Signs of a Cubic Zirconium School are resistance to scrutiny, a history of no research, descriptions that emphasize show rather than substance, and staff that are reluctant to talk to outsiders. The prognosis for surface-structural is moderately low; The prognosis for deep-structural change is low.

The Slightly Imperfect (SI) Diamond School

The Slightly Imperfect (SI) Diamond School is high in desire for actual change, low in desire for the appearance of change, and high in self-esteem. Its mood is one of denial. Its self-belief is that "If only we could get rid of 'X,' we'd be really good." "X" is a different thing in different schools, but it is the scapegoat for the schools' woes. The school is like a slightly imperfect diamond in that it has, from its own point of view, one not entirely unobvious flaw, which it would just as soon deny, if it could. Typical self-statements of the school are that, "We're pretty damn good, although we've got this 'X' to deal with," "If it weren't for 'X,' we'd be number one," and "We try to keep 'X' in line (ha-ha)."

Signs of the Slightly Imperfect Diamond School are praise of the system coupled with veiled digs at "X," deflection of probing questions about "X," attempts to deny the problem of "X," and generally favorable signs, but subtle hints that something is wrong. The prognosis for surface-structural change is moderately high, and for deep-structural change is moderately low. Indeed, if the problem of "X" can be successfully dealt with, the school will be in an excellent position to change.

The Lead School

The Lead School is high in desire for actual change, high in desire for appearance of change, but low in self-esteem. Its mood is one of superstitiousness. Its self-belief is that, "We need a quick way to turn lead into gold." The school has an almost alchemical or magical view that some quick fix will turn it into the kind of school it wants to be. Typical self-statements are that, "We'll give you a month to show what you can do"; "We can give you an hour per week"; "We need quick results here"; and "We want change, not research."

Signs of a Lead School are impatience, magical beliefs with respect to possibilities for change, lack of interest in understanding interventions, lack of understanding of programs, and an emphasis on doing, not planning. The prognosis for surface-structural change is moderately high, and that for deep-structural change is moderately low. If one can get the self-esteem of the school up, so that it is not forced to resort to superstitiousness, the prognosis can be excellent.

The Diamond in the Rough School

The Diamond in the Rough School is high in desire for actual change, desire for appearance of change, and self-esteem. Its mood is one of hopefulness. Its

self-belief is that "We've got the raw material here to be really great, and we're going to be." Examples of self-statements are that "We can make this work and we will," "We're on the way up," "You can help us be what we want to be," "There'll be problems but we can overcome them," and "We want to be great."

Signs of a Diamond in the Rough School are willingness to devote resources such as time and money to change, planfulness, accurate recognition of strengths and weaknesses, and receptiveness. The Diamond in the Rough School views itself in just this way, as a diamond that has a great deal of value, but needs to be shaped and formed. Often it will seek outsiders to help it do so. The prognosis for surface-structural change and for deep-structural change is high.

EXAMPLES OF THE VARIOUS KINDS OF SCHOOLS

To many of the readers of this chapter, at least some of the types of schools described will be familiar. Nevertheless, it may be helpful to illustrate some of the types of schools with examples of encounters that I've had in my past dealings with schools.

"Irontown" was a "Rusted-Iron" school. For starters, it was extremely difficult to find out where in the bureaucracy of the Irontown school district one should make a contact. Eventually we made a contact. The contact person, who was from the central office, did not seem to have much enthusiasm for meeting with us, but agreed to do so. Inexplicably, the meeting was later cancelled by a secretary. After yet another cancellation, the meeting finally took place. However, the official we contacted did not show up. The people who did show up seemed to be minor functionaries without much idea of (or interest in) how the system worked. We did make some progress in the meeting, however, and each party agreed to make some preparations for a future meeting. This meeting took place, but it was as though the first meeting had never happened: There had been no follow-through on the part of the school personnel. Moreover, they seemed to have only a foggy remembrance of the first meeting. This time, we tried specifying in writing what each team would do in order to facilitate interaction with the district. The next meeting was cancelled. When we eventually met again, once more, nothing had happened. We discontinued contact with the school district after this meeting.

Irontown is an urban school district in which most of the schools are inner-city schools with what might euphemistically be called 'hardened' teachers and students. The emphasis in the classrooms tends to be on discipline and on simply getting through the day. From what few observations we have had, there seems actually to be relatively little education going on. The district is generally considered to be one of the worst in its state, and our limited experience

with the administration led us to see no basis for believing that its standing would change.

“Granite Academy” was a parochial school that drew its students mostly from a blue-collar ethnic population. The school had a reputation for regimentation, but for giving students a fine education. I had no trouble making contact with the proper officials, and met with them promptly. The officials listened politely to my presentation for a suggested intervention. They seemed interested, and asked several questions. They told me that they would get back to me.

That they did the next day. Their decision was that they did not really want a research project in the school. They informed me that they had found what we had to say interesting, but that they felt that their programs did teach for thinking, and moreover, that their school wasn’t in need of fixing so there was no point in fixing it. It had been clear all along that they really did have high confidence in what they were doing, and that it would take hard persuasion to get them to change. Obviously, my persuading was not hard enough. In this particular case, the questioning attitude I sought to instill in the students was in direct conflict with the presuppositions of the school and many of the personnel in it. The emphasis in the school was on rote memorization of both secular and religious content, and the successful student was one who spit back the facts, or what would often seem to some to be opinions, rephrased as facts. It would have been difficult to succeed in this school, given its Granite culture.

Whereas Granite Academy saw itself as a bastion preserving traditional values, Opalville saw itself as at the forefront of modern education. This school had a high budget for educational materials, a high budget for in-service presentations, and some of the highest salaries in its state. A public school in a well-heeled town, Opalville was generally considered to be among the best schools in the state. They welcomed researchers as much as schools ever do—but almost exclusively for one-shot or two-shot research studies. They were interested in the research as part of their image, rather than as a vehicle for change. Moreover, the in-service pattern was one of one-shots. This month might be on school reform, the next on thinking skills, the next on drug education, the next on classroom management. The school paid good money to bring in nationally known consultants, but there seemed to be little follow-through. Again, one had the feeling that they were more concerned about the appearance of forward-looking in-service programs than they were about any real change. And indeed, the school did not appear to be changing much. Rather, it rested on a fine reputation and on putting a lot of money into education to give the appearance it wanted.

As a last example, consider Diamond City. Diamond City was comparable socioeconomically to Opalville, but their attitude toward school reform was very different. Whereas Opalville wanted only one-shot, brief research projects, Diamond City wanted only research projects that brought with them some real

chance of change in the schools. They encouraged us to get involved not only with the school administration and the teachers, but with the Board of Education and with parent groups as well. Our program became a sort of community effort, and we would occasionally get calls from interested parents about the program and what they could do to support it. The program lasted five weeks, and in terms of quantitative data, was among the most successful we've done. The combined support of administration, teachers, Board of Education, and parents made this the kind of experience any consultant would like to have.

Obviously, the success of an intervention will depend in large part on the quality of the intervention and on the skill with which it is executed. A good intervention stands a better chance of success in any school district, and a bad one is not likely to work in even the best districts. Nevertheless, it is useful to take into account the modifiability of the context in which one works. Interventions do not have an equal chance in every context in which they might be applied. It is useful for the intervener to understand the context into which he or she is going, and ultimately, for the school to understand as well what their attitude toward modifiability is.

Low-modifiable contexts can be made more highly modifiable, but to do so requires rather massive intervention in all aspects of the context, with the explicit goal of increasing modifiability. One cannot expect normal interventions to change the context markedly and durably—the context is more likely to change them. To change the modifiability of the context, the school needs to set up a task force with that particular goal in mind. The task force needs to identify what the cultural problems are in the school, to document these problems, to convince other people that these problems exist and that they are indeed problems, and then to work with the school as a whole for change. We need to realize that interventions can be of value in a less modifiable context, but one needs to be realistic with respect to what they are and are not likely to accomplish. Nevertheless, even small changes now may plant seeds for larger changes later on.

CONCLUSION

I have argued that intelligence can be understood fully only if it is conceptualized broadly, and if its style of use and the context within which it interacts is also understood. Consider some examples.

A person with the ability to cope well with novelty, but without a legislative style, may have the potential to do creative work, but will probably never fully exploit this potential. The person may be able to be creative, but won't like being that way. Even a person with high ability to cope with novelty and

a legislative style will probably be fully thwarted in a Rusted-Iron context, or any of a number of others. Or a person with high analytic ability and an executive style may thrive in an Opal context, but not in a Diamond-in-the-Rough one. A person with relatively low coping-with-novelty ability but who is characterized by a legislative style may be a frustrated poet or scientist, whereas a person with this same style but more coping-with-novelty ability may succeed as either a poet or a scientist. Of course, that success may be tempered by context. The individual may have the ability and style to be creative, but not be in a type of context that allows creativity to shine through.

Although the three issues discussed in this article can help us more fully understand adjustment and success in the everyday world, obviously they are not exhaustive. They do not fully account, say, for either creativity (see Sternberg & Lubart, 1991) or wisdom (see Sternberg, 1990a, 1990b). But they go much further toward understanding what happens in the world than do narrow psychometric theories of intelligence. One thing is for sure: If we want intelligently to understand and assess intelligence, we need to go beyond the conventional intelligence test.

REFERENCES

- Berry, J. W. (1972). Radical cultural relativism and the concept of intelligence. In L. J. Cronbach & P. Drenth (Eds.), *Mental tests and cultural adaptation*. The Hague, Holland: Mouton.
- Binet, A., & Simon, T. (1973). *Classics in psychology: The development of intelligence in children*. New York: Arno Press.
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, 7, 513–531.
- Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 1). Hillsdale, NJ: Erlbaum.
- Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In P. H. Mussen (Ed.), *Handbook of child psychology* (4th ed., Vol. 3, pp. 77–166). New York: Wiley.
- Butterfield, E. C., Wambold, C., & Belmont, J. M. (1973). On the theory and practice of improving short-term memory. *American Journal of Mental Deficiency*, 77, 654–669.
- Campione, J. C., & Brown, A. L. (1979). Towards a theory of intelligence: Contributions from research with retarded children. In R. J. Sternberg & D. K. Detterman (Eds.), *Human intelligence: Perspectives on its theory and measurement* (pp. 139–164). Norwood, NJ: Ablex.
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston, MA: Houghton Mifflin.

- Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic.
- Chi, M. T. H. (1978). Knowledge structure and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73–96). Hillsdale, NJ: Erlbaum.
- Cole, M., Gay, J., Glick, J., & Sharp, D. W. (1971). *The cultural context of learning and thinking*. New York: Basic Books.
- Cooper, L. A. (1982). Strategies for visual comparison and representation: Individual differences. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1, pp. 77–124). Hillsdale, NJ: Erlbaum.
- Davidson, J. E., & Sternberg, R. J. (1984). The role of insight in intellectual giftedness. *Gifted Child Quarterly*, 28, 58–64.
- Feuerstein, R. (1980). *Instrumental enrichment: An intervention program for cognitive modifiability*. Baltimore, MD: University Park Press.
- Flavell, J. H. (1977). *Cognitive development*. Englewood Cliffs, NJ: Prentice-Hall.
- Ford, M. E. (1986). A living systems conceptualization of social intelligence: Outcomes, processes, and developmental change. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 3, pp. 119–171). Hillsdale, NJ: Erlbaum.
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic.
- Gelman, R., & Baillargeon, R. (1983). A review of some Piagetian concepts. In P. Mussen (Series Ed.), J. Flavell & E. Markman (Vol. Eds.), *Handbook of child psychology*. New York: Wiley.
- Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.
- Gladwin, T. (1970). *East is a big bird*. Cambridge, MA: Harvard University Press.
- Gregorc, T. (1985). *Inside styles: Beyond the basics*. Maynard, MA: Gabriel Systems, Inc.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw Hill.
- Gustafsson, J. E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, 8, 279–303.
- Heath, S. S. (1983). *Ways with words*. New York: Cambridge University Press.
- Holland, J. L. (1973). *Making vocational choices: A theory of careers*. Englewood Cliffs, NJ: Prentice-Hall.
- Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, 75, 242–259.
- Horn, J. L., & Knapp, J. R. (1973). On the subjective character of the empirical base of Guilford's structure-of-intellect model. *Psychological Bulletin*, 80, 33–43.
- Hunt, E. B. (1978). Mechanics of verbal ability. *Psychological Review*, 85, 109–130.
- Hunt, E. B. (1980). Intelligence as an information-processing concept. *British Journal of Psychology*, 71, 449–474.
- Hunt, E. B., & Lansman, M. (1982). Individual differences in attention. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1). Hillsdale, NJ: Erlbaum.
- Irvine, S. H. (1979). The place of factor analysis in cross-cultural methodology and its contribution to cognitive theory. In L. Eckensberger & Y. Poortinga (Eds.), *Cross-cultural contributions to psychology*. Lisse, Netherlands: Swets & Zeitlinger.

- Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.
- Jensen, A. R. (1982). Reaction time and psychometric *g*. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 93–132). Berlin: Springer-Verlag.
- Jung, C. (1923). *Psychological types*. New York: Harcourt, Brace.
- Kearins, J. M. (1981). Visual spatial memory in Australian aboriginal children of desert regions. *Cognitive Psychology*, 2, 434–460.
- Keil, F. C. (1984). Transition mechanisms in cognitive development and the structure of knowledge. In R. J. Sternberg (Ed.), *Mechanisms of cognitive development* (pp. 81–99). New York: Freeman.
- Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.
- Laboratory of Comparative Human Cognition. (1982). Culture and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 642–719). New York: Cambridge University Press.
- MacLeod, C. M., Hunt, E. B., & Mathews, N. N. (1978). Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning and Verbal Behavior*, 17, 493–507.
- Mayer, R., & Greeno, J. G. (1972). Structural differences between learning outcomes produced by different instructional methods. *Journal of Educational Psychology*, 63, 165–173.
- Milgram, S. (1975). *Obedience to authority*. New York: Harper & Row.
- Mulholland, T. M., Pellegrino, J. W., & Glaser, R. (1980). Components of geometric analogy solution. *Cognitive Psychology*, 12, 252–284.
- Myers, I. B. (1980). *Gifts differing*. Palo Alto, CA: Consulting Psychologists Press.
- Myers, I. B., & McCauley, M. H. (1985). *Manual: A guide to the development and use of the Myers-Briggs type indicator*. Palo Alto, CA: Consulting Psychologists Press.
- Neisser, U. (1983). Components of intelligence or steps in routine procedures? *Cognition*, 15, 189–197.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Pellegrino, J. W., & Glaser, R. (1979). Cognitive correlates and components in the analysis of individual differences. In R. J. Sternberg & D. K. Detterman (Eds.), *Human intelligence: Perspectives on its theory and measurement* (pp. 61–88). Norwood, NJ: Ablex.
- Pellegrino, J. W., & Kail, R., Jr. (1982). Process analyses of spatial aptitude. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1, pp. 311–366). Hillsdale, NJ: Erlbaum.
- Piaget, J. (1972). *The psychology of intelligence*. Totowa, NJ: Littlefield, Adams.
- Raaheim, K. (1974). *Problem solving and intelligence*. Oslo, Norway: Universitetsforlaget.
- Renzulli, J. S., & Smith, L. H. (1978). *Learning styles inventory*. Mansfield Center, CT: Creative Learning Press.
- Resnick, L. B., & Glaser, R. (1976). Problem solving and intelligence. In L. B. Resnick (Ed.), *The nature of intelligence*. Hillsdale, NJ: Erlbaum.
- Schank, R. C. (1980). How much intelligence is there in artificial intelligence? *Intelligence*, 4, 1–14.
- Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Erlbaum.

- Siegler, R. S. (1986). Unities of thinking across domains. In M. Perlmutter (Ed.), *The development of intelligence: The 1984 Minnesota symposium on child psychology*. Hillsdale, NJ: Erlbaum.
- Simon, H. A., & Reed, S. (1976). Modeling strategy shifts in a problem-solving task. *Cognitive Psychology*, 8, 86–97.
- Snow, R. E. (1979). Theory and method for research on aptitude processes. In R. J. Sternberg & D. K. Detterman (Eds.), *Human intelligence: Perspectives on its theory and measurement* (pp. 105–138). Norwood, NJ: Ablex.
- Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 47–104). Hillsdale, NJ: Erlbaum.
- Snow, R. E., & Lohman, D. F. (1984). Toward a theory of cognitive aptitude for learning from instruction. *Journal of Educational Psychology*, 75, 347–376.
- Spearman, C. (1923). *The nature of 'intelligence' and the principles of cognition*. London: Macmillan.
- Spearman, C. (1927). *The abilities of man*. New York: Macmillan.
- Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Erlbaum.
- Sternberg, R. J. (1979). The nature of mental abilities. *American Psychologist*, 34, 214–230.
- Sternberg, R. J. (1980). Sketch of a componential subtheory of human intelligence. *Behavioral and Brain Sciences*, 3, 573–584.
- Sternberg, R. J. (1981). Intelligence and nonentrenchment. *Journal of Educational Psychology*, 73, 1–16.
- Sternberg, R. J. (1982a). Natural, unnatural, and supernatural concepts. *Cognitive Psychology*, 14, 451–458.
- Sternberg, R. J. (1982b). Reasoning, problem solving, and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 225–307). New York: Cambridge University Press.
- Sternberg, R. J. (1983). Components of human intelligence. *Cognition*, 15, 1–48.
- Sternberg, R. J. (1985a). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.
- Sternberg, R. J. (1985b). Human intelligence: The model is the message. *Science*, 230, 1111–1118.
- Sternberg, R. J. (1986). *Intelligence applied: Understanding and increasing your intellectual skills*. San Diego: Harcourt, Brace, Jovanovich.
- Sternberg, R. J. (1987). The psychology of verbal comprehension. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 3, pp. 97–151). Hillsdale, NJ: Erlbaum.
- Sternberg, R. J. (1988). Mental self-government: A theory of intellectual styles and their development. *Human Development*, 31, 197–224.
- Sternberg, R. J. (1990a). *Metaphors of mind: Conceptions of the nature of intelligence*. New York: Cambridge University Press.
- Sternberg, R. J. (Ed.). (1990b). *Wisdom: Its nature, origins, and development*. New York: Cambridge University Press.
- Sternberg, R. J., & Davidson, J. E. (1982). The mind of the puzzler. *Psychology Today*, 16, June 37–44.

- 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* (pp. 231–254). Berlin: Springer-Verlag.
- Sternberg, R. J., & Ketron, J. L. (1982). Selection and implementation of strategies in reasoning by analogy. *Journal of Educational Psychology, 74*, 399–413.
- Sternberg, R. J., & Lubart, T. I. (1991). An investment theory of creativity and its development. *Human Development, 34*, 1–31.
- Sternberg, R. J., & Powell, J. S. (1983). Comprehending verbal comprehension. *American Psychologist, 38*, 878–893.
- Sternberg, R. J., & Rifkin, B. (1979). The development of analogical reasoning processes. *Journal of Experimental Child Psychology, 27*, 195–232.
- Sternberg, R. J., & Suben, J. (1986). The socialization of intelligence. In M. Perlmutter (Ed.), *Perspectives on intellectual development (Vol. 19): Minnesota symposia on child psychology* (pp. 201–235). Hillsdale, NJ: Erlbaum.
- Sternberg, R. J., & Wagner, R. K. (Eds.). (1986). *Practical intelligence: Nature and origins of competence in the everyday world*. New York: Cambridge University Press.
- Sternberg, R. J., & Weil, E. M. (1980). An aptitude-strategy interaction in linear syllogistic reasoning. *Journal of Educational Psychology, 72*, 226–234.
- Strong, E. K., Jr., Campbell, D. P., & Hansen, J. C. (1985). *Strong-Campbell Interest Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Super, C., & Harkness, S. (1980). The infant's niche in rural Kenya and metropolitan America. In L. L. Adler (Ed.), *Issues in cross-cultural research*. New York: Academic.
- Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.
- Vernon, P. E. (1971). *The structure of human abilities*. London: Methuen.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wagner, R. K., & Sternberg, R. J. (1987). Executive control in reading comprehension. In B. K. Britton & S. M. Glynn (Eds.), *Executive control processes in reading* (pp. 1–21). Hillsdale, NJ: Erlbaum.
- Wigdor, A. K., & Garner, W. R. (Eds.). (1982). *Ability testing: Uses, consequences, and controversies*. Washington, DC: National Academy Press.

Author Index

A

- Abelson, R.P., 369, 405
Ackerman, P.L., 7, 11, 12, 13, 17, 22, 23, 23,
24, 25, 320, 322, 340, 344, 350, 354,
357
Adams, J.A., 11, 23
Ahern, F.M., 86, 114
Aiello, P., 296, 305
Alberch, P., 79, 81
Alderton, D.L., 272, 283, 315, 348, 356
Aleksander, I., 40, 40
Alexander, W.P., 120, 142, 160, 197
Allen, L., 4, 24
Anderson, J.E., 218, 230
Anderson, J.R., 309, 310, 342, 342, 345, 347,
354
Anderson, N.S., 340, 356
Atkin, R., 217, 230
Atkinson, J.W., 16, 23
Augustin, J.W., 336, 360

B

- Bachor, D.G., 52, 53, 62
Baddeley, A.D., 245, 254, 336, 351, 354
Bahrick, H.P., 247, 254
Baillargeon, R., 381, 404
Baker, C.T., 22, 25
Baltes, P.B., 175, 178, 179, 197, 202
Bamberger, J., 290, 302
Barclay, C., 296, 305
Barrett, G.V., 67, 81
Barrett, P., 130, 131, 132, 133, 134, 137, 142

- Barron, F., 141, 142
Bayley, N., 4, 23
Bechtoldt, H.P., 52, 62
Belmont, J.M., 364, 403
Bergeman, C.S., 67, 83
Berger, M., 128, 142, 143
Bernstein, M., 18, 26, 121, 148
Berry, J.W., 381, 403
Biederman, I., 162, 197
Binet, A., 4, 23, 88, 112, 125, 143, 157, 197,
378, 403
Birren, J.E., 184, 198
Blackwood, W., 195, 197
Bloom, A., 31, 40
Blythe, T., 301, 305
Bobrow, D.G., 13, 26
Bodmer, 73, 81
Bock, R.D., 336, 360
Bordelon, V.P., 319, 354
Boring, E.G., 118, 143, 157, 298, 302
Boruch, R.F., 342, 360
Botzum, W.A., 52, 62
Bouchard, T.J., 31, 40, 251, 254
Bourne, L.R., 195, 197
Bower, G.H., 352, 356
Bramble, W.J., 172, 184, 200
Brand, C.R., 30, 31, 32, 33, 34, 35, 39, 40, 41
Bransford, J.D., 366, 403
Bray, R., 217, 230
Bregman, E.O., 10, 27, 121, 149
Brierley, H., 128, 143
Brigham, C.C., 160, 198

- Broadbent, D.E., 162, 198
 Bronfenbrenner, U., 72, 75, 78, 80, 81, 384, 403
 Brookings, J., 315, 320, 350, 354
 Brooks, L.R., 245, 254
 Brookshire, R.G., 184, 199
 Brown, A.L., 295, 301, 302, 303, 364, 366, 403
 Brown, J.S., 295, 301, 303
 Brown, W., 160, 198
 Browning, J.L., 268, 283
 Bruner, J., 30, 41
 Buchsbaum, M.S., 268, 283
 Buktenica, N.A., 128, 143
 Burt, C., 122, 125, 126, 143, 159, 166, 167, 198
 Butcher, H.J., 142, 143
 Butler, J.M., 3, 5, 6, 22, 23, 24
 Butterfield, E.C., 364, 403
- C**
- Callaway, E., 133, 143
 Cameron, N., 141, 143
 Campbell, D.P., 407
 Campbell, D.T., 131, 143, 174, 198
 Campione, J.C., 364, 403
 Card, S.K., 310, 340, 355
 Cardon, L.R., 48, 62, 216, 230
 Cariglia-Bull, T., 289, 304
 Carlson, J.S., 126, 143
 Carlson, R.A., 338, 355
 Carpenter, P.A., 337, 355
 Carretta, T.R., 319, 355
 Carroll, J.B., 43, 44, 48, 62, 63, 68, 81, 110, 112, 123, 143, 155, 161, 167, 171, 172, 173, 198, 255, 266, 275, 276, 282, 315, 336, 355
 Carsrud, A.L., 15, 24
 Caruso, D.R., 70, 82, 108, 109, 113
 Caryl, P.G., 31, 41, 132, 133, 137, 143
 Cattell, J.M., 142, 143
 Cattell, R.B., 4, 20, 23, 24, 59, 60, 63, 86, 112, 120, 142, 143, 146, 160, 163, 167, 168, 172, 173, 175, 183, 188, 198, 199, 200, 309, 355, 372, 379, 403
 Cavalli-Sforza, 73, 81
 Ceci, S.J., 68, 69, 71, 72, 75, 76, 78, 79, 80, 81, 82, 287, 303
 Cerella, J., 247, 255
 Chadwick, S.B., 138, 146
 Chaiken, S., 312, 338, 342, 343, 353, 355
 Chase, W.G., 368, 404
 Chastain, R.L., 359
 Cheng, P.W., 244, 255
 Cheung, P.C., 19, 26
 Chi, M.T.H., 368, 404
 Chipuer, H.M., 187, 198
 Christal, R.E., 12, 13, 25, 311, 312, 315, 316, 325, 326, 330, 337, 338, 339, 345, 355, 357
 Claridge, G., 139, 149
 Clark, L.A., 324, 360
 Cobb, M.V., 10, 27, 121, 149, 287, 305
 Cohen, J., 325, 355
 Cohen, P., 325, 355
 Cole, M., 305, 374, 404
 Cole, R.E., 86, 114
 Collins, A., 295, 301, 303
 Commons, M., 155, 198
 Conners, F.A., 70, 82, 108, 109, 111, 112, 113
 Conway, B.E., 18, 26, 121, 148
 Cookson, D., 142, 145
 Cooper, L.A., 365, 404
 Coresellis, J.A., 195, 197
 Costa, P.T., Jr., 38, 41
 Coulson, J., 119, 147
 Cox, G.W., 160, 198
 Crano, W.D., 120, 148
 Crawford, J.D., 284, 202
 Crinella, F.M., 89, 110, 114
 Cronbach, L.J., 3, 5, 10, 16, 20, 23, 154, 174, 198, 322, 323, 329, 330, 355
 Crutchfield, R.S., 141, 143
 Csikszentmihalyi, M., 290, 303
 Cunitz, A.R., 162, 199
 Cunningham, W.R., 184, 198
 Curran, L.T., 336, 360
- D**
- Daneman, M., 337, 355
 Daniel, M.H., 39, 41, 70, 82, 108, 113, 137, 144, 248, 250, 255
 Daum, I., 132, 142
 Davey, T.C., 217, 231
 Davidson, J.E., 245, 255, 318, 347, 355, 370, 404, 406
 Deary, J.J., 31, 41, 70, 82, 132, 133, 137, 143
 DeFries, J.C., 48, 62, 121, 147, 187, 190, 199, 201, 216, 230
 Dekaban, K.A., 253, 255

- Dennis, I., 348, 356
 Depinet, R.L., 67, 81
 Detterman, D.K., 39, 41, 69, 70, 72, 82, 88, 89, 108, 109, 112, 112, 113, 118, 127, 137, 144, 148, 158, 161, 199, 202, 248, 250, 255, 258, 275, 282, 284
 De Vries, R., 123, 143
 Division, M., 217, 230
 Dixit, S., 322, 355
 Dominowski, R.L., 195, 197
 Donaldson, G., 162, 172, 176, 179, 184, 187, 191, 200
 Dreger, R.M., 17, 24
 Durea, M.A., 5, 24
 Dweck, C.S., 19, 24
- E**
 Earles, J.A., 266, 284, 330, 358
 Eaves, L.J., 32, 41, 191, 199
 Eckstrom, R.B., 168, 169, 199
 Egan, V., 31, 32, 33, 34, 35, 41
 Ekstrand, B.R., 195, 197
 Eliot, J., 245, 255
 El Koussy, A.A.H., 160, 199
 Elliot, R.N., 160, 201
 Ellis, N.R., 88, 113
 Engle, R.W., 337, 356, 360
 Engstrom, R., 162, 184, 187, 191, 200
 Eysenck, H.J., 31, 41, 67, 82, 89, 95, 113, 117, 118, 120, 121, 122, 124, 125, 126, 127, 128, 130, 131, 132, 133, 134, 136, 137, 138, 139, 142, 142, 144, 145, 160, 161, 191, 199, 248, 255, 287, 255
 Eysenck, S.B.G., 138, 139, 145
- F**
 Fagan, J.F., 126, 145
 Fairbank, B.A., Jr., 336, 340, 356, 360
 Falconer, D.S., 189, 199
 Faust, M.E., 318, 356
 Feldman, D.H., 288, 290, 299, 303
 Ferguson, G.A., 10, 24
 Ferrara, R.A., 366, 403
 Feuerstein, R., 391, 404
 Finkbeiner, C.T., 50, 63
 Fischbein, S., 78, 82
 Fischer, B., 80, 83, 280, 283
 Fiske, D.W., 3, 5, 6, 22, 24, 174, 198
 Fitts, P.M., 11, 24
 Flavell, J.H., 364, 404
 Flory, C.D., 4, 24
 Flynn, J.R., 31, 41, 219, 230, 278, 282
 Fodor, J., 289, 303
 Ford, M.E., 384, 404
 Fowler, H.V., 119, 147
 Frearson, W., 126, 130, 131, 145
 French, J.W., 46, 63, 168, 169, 199, 341, 356
 Frick, R.W., 315, 348, 356
 Fulker, D.W., 48, 62, 216, 230
 Furneaux, D., 128, 145
- G**
 Gagne, R.M., 309, 356
 Gallistel, C.R., 366, 404
 Galton, F., 86, 113, 125, 145
 Gardner, H., 121, 145, 233, 255, 295, 299, 300, 301, 303, 304, 367, 384, 404
 Gardner, M.K., 378, 384, 407
 Garner, W.R., 361, 407
 Garnett, J.C.M., 100, 113
 Gay, J., 374, 404
 Gelman, R., 366, 381, 404
 Gergen, K.J., 30, 41
 Gernsbacher, M.A., 318, 356
 Getzels, J.W., 139, 145
 Gillet, A.H., 336, 360
 Gilliland, K., 5, 26
 Gladwin, T., 374, 404
 Glanzer, M., 162, 199
 Glaser, R., 364, 367, 405
 Glick, J., 374, 404
 Gluck, M.A., 352, 356
 Glymour, C., 335, 336, 356
 Goff, M., 23, 24
 Goldsmith, H.H., 190, 192, 201
 Goodnow, J.J., 295, 298, 303
 Gordon, I.E., 34, 26
 Gordon, L.V., 5, 24
 Gordon, R.A., 41
 Gottfredson, L.S., 67, 82
 Götz, K., 140, 145
 Götz, K.O., 140, 145
 Gould, P., 348, 356
 Gould, S.J., 287, 303
 Green, P.B., 79, 84
 Greeno, J.G., 301, 303, 368, 405
 Gregorc, T., 394, 404
 Gregory, R.L., 40, 41
 Guerrero, C., Jr., 336, 360

- Guilford, J.P., 8, 24, 98, 107, 113, 121, 145,
168, 199, 233, 255, 287, 303, 329, 330,
356, 379, 404
- Gustafsson, J.-E., 60, 63, 121, 122, 145, 149,
162, 163, 167, 172, 199, 325, 330, 356,
379, 404
- Guttman, L., 122, 145, 210, 230

H

- Haier, R.J., 268, 282
- Hakstian, A.R., 59, 60, 63, 167, 172, 173, 199
- Hale, S., 251, 255
- Hall, L.K., 247, 254
- Hammer, M., 138, 145
- Hansen, J.C., 407
- Harkness, S., 375, 407
- Harman, H.H., 166, 199
- Harman, M.H., 168, 169, 199
- Harmquist, K., 220, 222, 230
- Hart, B., 90, 113
- Harter, S., 35, 41
- Hartigan, J., 228, 231
- Hatano, G., 296, 303
- Hayslip, B., 184, 199
- Hazlett, E., 268, 282
- Heath, S.S., 375, 404
- Hellewell, T.B., 83
- Helmreich, R.L., 15, 24
- Hemmelgarn, T.E., 127, 146
- Hendrickson, A.E., 67, 82, 133, 136, 146
- Hendrickson, D.E., 67, 82, 133, 146
- Henri, V., 88, 112
- Henry, R.A., 14, 24, 25
- Herrnstein, R.J., 75, 82, 188, 199, 250, 255
- Hertzog, C., 173, 199
- Herzberger, S., 217, 230
- Heston, L.L., 138, 146
- Hewlett, J.H., 141, 147
- Hinton, G.E., 240, 256
- Hoennmann, H.-J., 121, 146
- Hoepfner, R., 121, 145
- Holland, J.L., 395, 404
- Holzinger, K.J., 166, 199
- Holyoak, K.J., 244, 255
- Honzik, M.P., 4, 24
- Horn, J.L., 4, 21, 24, 59, 63, 70, 82, 98, 115,
120, 121, 123, 126, 146, 148, 149, 153,
161, 162, 163, 167, 171, 172, 173, 175,
176, 177, 179, 183, 184, 187, 188, 191,
192, 194, 198, 199, 200, 201, 220, 231,
246, 255, 329, 356, 379, 404

- House, B.J., 88, 115
- Hoyle, F., 121, 146
- Hudson, L., 140, 146
- Hulin, C.L., 14, 24, 25
- Hultsch, D.F., 201
- Humphreys, L.G., 4, 14, 15, 18, 25, 66, 75, 82,
99, 113, 124, 146, 156, 164, 171, 172,
201, 208, 210, 211, 212, 216, 217, 222,
224, 228, 230, 231, 356
- Humphreys, M.S., 5, 17, 25, 26
- Hunt, E.B., 46, 63, 89, 113, 240, 245, 253, 255,
315, 320, 321, 348, 350, 351, 356, 360,
365, 367, 371, 376, 404, 405
- Hunter, J.E., 67, 83, 223, 231, 232, 267,
284
- Husen, T., 4, 25

I

- Irvine, S.H., 348, 356, 381, 404
- Iverson, L.L., 194, 201

J

- Jackson, A., 301, 305
- Jackson, D., 348, 359
- Jackson, P.W., 139, 145
- Jäger, A.O., 121, 122, 146
- Jarvik, L.F., 138, 146
- Jaschek, L., 121, 146
- Jaschek, M., 121, 146
- Jencks, C., 287, 303
- Jensen, A.R., 66, 67, 83, 89, 92, 93, 95, 110,
113, 114, 122, 126, 146, 147, 160, 161,
164, 193, 201, 220, 231, 258, 264, 266,
268, 269, 270, 271, 272, 273, 274, 275,
276, 278, 279, 280, 283, 284, 287, 303,
304, 339, 356, 366, 367, 372, 380, 405
- Johnson, E.S., 44, 63
- Johnson, M.B., 174, 203
- Johnson, R.C., 86, 114, 122, 147
- Johnson-Laird, P.N., 246, 256, 310, 357
- Jones, L., 148
- Jung, C., 394, 405
- Just, M.A., 337, 355

K

- Kagan, J., 25
- Kahneman, D., 7, 8, 25
- Kail, R., Jr., 367, 405
- Kanfer, R., 11, 13, 17, 19, 22, 25, 322, 357
- Kantor, J.E., 319, 354
- Karlsson, J.L., 138, 147

Kaufman, A.S., 138, 147, 183, 201
 Kaufman, H., 10, 27
 Kaufman-Packer, J.L., 183, 201
 Kearins, J.M., 374, 405
 Kehle, T.J., 127, 146
 Keil, F.C., 368, 405
 Kelley, T.L., 160, 201
 Kelly, K., 335, 336, 356
 Keppel, G., 359
 Ketron, J.L., 18, 26, 121, 148, 365, 407
 Khoo, B.H., 338, 355
 Knapp, J.R., 121, 146, 171, 200, 329, 356, 379, 404
 Koenig, O., 289, 299, 304
 Kogan, N., 139, 149
 Kohlberg, L., 123, 143
 Körkel, J., 71, 83
 Kornhaber, M., 295, 304
 Kosslyn, S.M., 289, 299, 304, 365, 405
 Kranzler, J.H., 110, 114, 272, 275, 276, 278, 283
 Kraska, K., 22, 25
 Krechevsky, M., 295, 299, 300, 301, 303, 304
 Kuhl, J., 17, 22, 25
 Kuhn, T.S., 117, 127, 147, 153, 175, 201
 Kuse, A.R., 190, 199
 Kyllonen, P.C., 12, 13, 23, 25, 122, 148, 248, 256, 311, 312, 313, 315, 316, 321, 325, 326, 330, 337, 338, 339, 340, 341, 343, 344, 345, 347, 372, 406

L

Lampert, M., 301, 304
 Langley, P., 310, 347, 357
 Lansman, M., 240, 255, 365, 404
 Larson, G.E., 272, 283
 Lashley, K.S., 93, 110, 114
 Last, K.A., 191, 199
 Lauer, J., 19, 26
 Lave, J., 295, 304
 Leggett, E.L., 19, 24
 Legree, P.J., 70, 82, 108, 109, 113
 Lehl, S., 80, 83, 280, 283
 Leiman, J.M., 50, 63
 Lenzenweger, M., 72, 75, 78, 80, 81
 Lewis, C.T., 119, 120, 147
 Lewis, J., 89, 113
 Lickliter, R., 83
 Liker, J., 71, 72, 76, 82
 Lin, P.C., 14, 15, 18, 25
 Lindsey, G., 187, 201

Linn, R., 228, 231
 Little, W., 119, 147
 Loehlin, J.C., 187, 201
 Loevinger, J., 121, 147
 Logan, G.D., 345, 357
 Lohman, D., 66, 83, 122, 147, 148, 334, 358, 372, 406
 Lubart, T.I., 403, 407
 Lubinski, D., 216, 228, 232
 Lunneborg, C., 89, 113
 Luria, A.R., 247, 255
 Lykken, D.T., 31, 40, 251, 254
 Lynn, R., 16, 26
 Lytton, H., 141, 147

M

MacFarlane, J., 4, 24
 MacIntyre, A., 30, 41
 Mackie, J., 100, 114
 MacKinnon, D.W., 141, 147
 MacLeod, C.M., 245, 253, 255, 318, 358
 MacLeod, C.W., 46, 63
 Malkus, U., 399, 304
 Malmi, R.A., 342, 360
 March, M., 325, 355
 Margaret, A., 141, 143
 Marshalek, B., 66, 83, 122, 147, 148, 248, 256, 316, 359, 372, 406
 Martin, N.G., 31, 32, 41, 191, 199
 Matarazzo, J.D., 180, 201
 Mathews, N.N., 46, 63, 365, 376, 405
 Maxwell, A.E., 100, 114, 121, 147
 Mayer, J.D., 82, 108, 109, 113
 Mayer, R., 368, 405
 McArdle, J.J., 177, 179, 180, 190, 191, 192, 200, 201
 McCaulley, M.H., 405
 McClearn, G.E., 68, 80, 83, 86, 114, 121, 147, 187, 190, 201
 McClelland, J.L., 240, 256
 McCrae, R.R., 38, 41
 McGrew, W.C., 289, 304
 McGue, M., 31, 40, 251, 254
 McLean, J.E., 183, 201
 McNeil, T.F., 138, 147
 McNemar, Q., 329, 358
 Meade, A., 44, 63
 Metcalfe, J., 348, 358
 Meyerson, J., 251, 255
 Milgram, S., 377, 405
 Minstrell, J., 253, 255

- Moran, T.P., 310, 340, 355
 Mori, M., 132, 149
 Mulholland, T.M., 367, 405
 Munro, E., 89, 113
 Mycielska, K., 6, 26
 Myers, I.B., 38, 41, 405
- N**
- Nagoshi, C.T., 86, 114, 147
 Neiderhiser, J.M., 67, 83
 Neiss, R., 31, 42
 Neisser, V., 121, 147, 297, 304, 367, 405
 Nelson, V.L., 22, 25
 Nesselroade, J., 68, 80, 83
 Nettlebeck, T., 147
 Newell, A., 238, 242, 255, 310, 340, 347, 358, 366, 405
 Newman, S.E., 301, 303
 Nicholls, J.G., 19, 26
 Nichols, R., 189, 190, 201
 Nisbett, R.E., 244, 253, 255
 Nolen, S.B., 19, 26
 Noon, S.L., 14, 25
 Norman, D.A., 13, 26, 358
 Nottebohm, F., 289, 304
 Nuechterlein, K.H., 268, 282
- O**
- Okagaki, L., 287, 301, 303, 304, 305
 Oliver, L.M., 244, 255
 Olson, D.R., 296, 304
 Ones, D.S., 267, 284
- P**
- Pagliari, H.C., 31, 41, 70, 82
 Palincsar, A.S., 301, 303
 Palmer, J., 245, 255, 318, 358
 Parasuraman, R., 10, 26
 Parfit, D., 30, 42
 Park, J., 268, 282
 Park, R.K., 222, 231
 Parsons, C.K., 124, 146, 222, 231
 Patashnick, M., 19, 26
 Patterson, D.G., 160, 201
 Payne, R.W., 141, 147
 Pea, R., 295, 305
 Pedersen, N.L., 68, 80, 83
 Pellegrino, J.W., 315, 320, 323, 348, 350, 351, 358, 360, 367, 405
 Pena, M.C., 312, 337, 346, 358, 359
 Perkins, D., 289, 295, 300, 304
- Persanyi, M., 70, 82
 Piaget, J., 292, 304, 380, 405
 Pieters, J.O.P.M., 341, 358
 Plomin, R., 48, 62, 66, 67, 68, 80, 83, 121, 147, 187, 190, 201, 216, 230, 252, 256
 Poon, L., 251, 255
 Popper, K., 148
 Porteus, S.D., 165, 201
 Posner, M.I., 11, 24, 308, 358
 Powell, J.S., 368, 407
 Pressley, M., 289, 304
 Price, L.A., 168, 199
 Prohovnik, I., 195, 201
- Q**
- Quereshi, M.Y., 5, 26
- R**
- Raaheim, K., 370, 405
 Rauschenberger, J., 67, 83
 Raven, J.C., 165, 201
 Raz, N., 126, 148
 Reason, J.T., 6, 26, 358
 Ree, M.J., 266, 284, 330, 358
 Reed, S., 366, 406
 Reed, T.E., 271, 272, 273, 274, 275, 283, 284
 Renzulli, J.S., 395, 405
 Resnick, L., 295, 304, 364, 405
 Revelle, W., 5, 16, 17, 18, 25, 26, 31, 42, 324, 358
 Revensdorff, D., 123, 148
 Reynolds, C.R., 183, 201
 Rifkin, B., 364, 366, 407
 Rimoldi, H.J.A., 160, 201
 Robinson, R., 290, 303
 Roe, A., 142, 148
 Roff, M., 218, 232
 Rogoff, B., 295, 296, 304
 Rossman, B.B., 172, 201
 Rothman, S., 118, 148
 Rovine, M.J., 187, 198
 Ruiz, A., 71, 72, 82
 Rumelhart, D.E., 240, 256
- S**
- Salomon, G., 289, 304
 Salter, W., 93, 94, 114
 Sattler, J.M., 5, 26, 387, 304
 Sawin, L.L., 15, 24
 Scarr, S., 67, 68, 83
 Schacter, D.L., 352, 359

- Schafer, E.W.P., 67, 83, 89, 114
 Schaie, K.W., 175, 178, 179, 183, 197, 202
 Shank, R.C., 368, 369, 405
 Scharfstein, B.A., 30, 42
 Scheines, R., 335, 336, 356
 Schmid, J., 50, 63
 Schmidt, F.L., 67, 83, 120, 148, 223, 232, 267, 284
 Schneider, W., 71, 83, 320, 338, 350, 354, 355
 Schoenemann, P.H., 188, 202
 Schoenfeld, A.H., 301, 304
 Schutz, R.E., 45, 46, 63
 Scribner, S., 247, 256, 389, 305
 Segal, N.L., 31, 40, 251, 254
 Selzer, U., 217, 230
 Shallice, T., 245, 256
 Sharp, D.W., 374, 404
 Sharp, S.E., 87, 114
 Shields, J., 120, 148
 Short, C., 119, 120, 147
 Shotter, J., 30, 41
 Shuell, T.J., 342, 359
 Shute, V.J., 312, 326, 339, 341, 342, 343, 344, 345, 357, 359, 360
 Siegel, B.V., 268, 282
 Siegler, R.S., 83, 365, 406
 Simon, H.A., 238, 242, 255, 366, 368, 378, 404, 405, 406
 Simon, L., 5, 26
 Simon, T., 4, 23, 88, 112, 157, 197
 Simonton, D.K., 131, 148
 Sinha, S.N., 268, 283
 Smith, G.J., 84, 251, 255
 Smith, L.H., 395, 405
 Sneath, P., 121, 124, 148
 Snow, R.E., 10, 16, 24, 26, 83, 122, 147, 148, 248, 256, 316, 329, 348, 355, 359, 366, 372, 378, 380, 405
 Snyder, B.L., 289, 304
 Snyderman, M., 118, 148
 Sokal, R., 121, 124, 148
 Sontag, L.W., 22, 25
 Soule, C.S., 344, 359
 Spear, L.C., 89, 114
 Spear, N.E., 84
 Spearman, C., 7, 26, 66, 84, 90, 113, 114, 122, 126, 138, 148, 156, 158, 159, 202, 249, 305, 308, 309, 359, 378, 406
 Sperling, G., 161, 162, 202
 Spirtes, P., 335, 336, 356
 Spuhler, J.N., 187, 201
 Squire, L.R., 347, 359
 Standskov, H.H., 190, 202
 Stankov, L., 126, 148, 184, 187, 188, 200, 271, 284
 Stephens, D.L., 312, 313, 339, 341, 343, 344, 357
 Stern, W., 157, 202
 Sternberg, R.J., 6, 18, 26, 27, 34, 42, 73, 84, 89, 93, 94, 114, 118, 121, 124, 125, 148, 158, 161, 174, 202, 238, 256, 258, 284, 287, 301, 303, 304, 305, 347, 355, 359, 361, 364, 365, 366, 367, 368, 369, 370, 371, 372, 375, 379, 380, 381, 382, 834, 392, 403, 404, 406, 407
 Sticht, T.G., 318, 359
 Stigler, J.W., 296, 305
 Strong, E.K., Jr., 407
 Suben, J., 407
 Subtelny, S., 79, 84
 Super, C., 375, 407
 Suppe, F., 117, 148
- T**
- Taber, T., 18, 25, 224, 231
 Taylor, R., 70, 82, 108, 109, 113
 Tellegen, A., 31, 40, 252, 254, 324, 360
 Teresi, J., 325, 355
 Terman, L.M., 157, 202
 Theye, F., 26
 Thompson, G.A., 164, 202
 Thompson, R., 110, 114
 Thomson, G.H., 99, 114, 121, 149
 Thorndike, E.L., 4, 10, 27, 99, 114, 121, 149, 289, 305, 322, 359
 Thurstone, L.L., 133, 149, 167, 177, 190, 202, 237, 256, 287, 292, 305, 309, 316, 341, 359, 379, 407
 Thurstone, T.R., 133, 149
 Tirre, W.C., 312, 337, 339, 340, 342, 345, 346, 356, 357, 358, 359
 Torrance, E.P., 140, 149
 Tucker, L.R., 50, 63
 Tuddenham, R.D., 86, 115, 125, 149, 219, 232
 Turner, M.L., 360
- U**
- Underwood, B.J., 310, 342, 360
 Undheim, J.O., 98, 115, 121, 149, 162, 167, 171, 172, 202, 203

V

- Valez, C., 325, 355
 Vandeberg, S.G., 1990, 199
 Van der Ven, A.H.G.S., 341, 358
 Varner, K.R., 318, 356
 Vernon, P.A., 132, 149, 271, 284
 Vernon, P.E., 8, 27, 121, 149, 372, 379, 407
 Vygotsky, L.S., 370, 407

W

- Wagner, R.K., 6, 27, 124, 148, 287, 305, 361, 366, 407
 Wagstaff, D., 251, 255
 Walker, C.H., 71, 84
 Wallach, M.A., 139, 149
 Walters, J., 301, 305
 Wambold, C., 364, 403
 Wankowsky, J.A., 142, 149
 Warrington, E.K., 289, 305
 Wason, P.C., 246, 256
 Watson, D., 324, 360
 Webb, N., 122, 148
 Wechsler, D., 4, 27
 Weil, E.M., 365, 407
 Weinert, F.E., 71, 83
 Weiss, V., 80, 84
 Weng, L.-J., 92, 95, 113, 147, 264, 283
 Whipple, G.M., 87, 115
 White, N., 301, 305
 White, P.O., 128, 149
 Whitehead, A.N., 8, 27
 Wickens, C.D., 8, 27, 245, 256, 320, 328, 350, 351, 354, 360

- Widaman, K.F., 126, 143
 Wigdor, A.K., 361, 407
 Wigdor, S., 228, 231
 Willerman, L., 126, 148
 Willoughby, R.R., 160, 203
 Wissler, C., 87, 115, 125, 149
 Wolman, B.B., 125, 149
 Woltz, D.J., 12, 25, 27, 312, 313, 326, 339, 342, 343, 344, 346, 353, 357, 360
 Woodcock, R.W., 161, 172, 174, 203
 Woodrow, H., 10, 27
 Woodworth, R.S., 289, 305
 Woody, E., 139, 149
 Woodyard, E., 10, 27, 121, 149, 287, 305
 Wothke, W., 316, 336, 360
 Wright, D.A., 348, 356
 Wright, G.H., 118, 149
 Wu, J.C., 268, 282

Y

- Yalow, E., 16, 26, 122, 148
 Yama, M., 126, 148
 Yaure, R.G., 338, 355
 Yeats, W.B., 39, 42
 Yee, P., 315, 320, 323, 348, 350, 358, 360
 Young, P.A., 191, 199
 Yu, J., 110, 114
 Yuen, S., 86, 114

Z

- Zeaman, D., 7, 10, 27, 88, 115
 Zinbarg, R.Z., 31, 42
 Zubin, J., 138, 145

Subject Index

- A**
Ability, 43–62, 151–197, 259, 308; *see also*
 Cognitive capability
 and accuracy, 321–322, 340–341
 and automatizing information processing,
 372–373
 -as-capacity, 4–5, 118, 119, 120, 121, 125,
 164, 240–242
 and factor analysis, 44ff., 69ff.
 and general intelligence, 94–112
 and genetics, 68–80, 187–195
 and intelligence tests, 93–94
 -as-maximal, 19
 measurement of, 157–158, 160–162,
 248–250, 260–266
 a model for, 311–327
 and mental processes, 88–90, 234ff.,
 259–260, 269–282
 and methodology, 44ff.
 number of, 155–156
 and novelty, 370–372
 and problem solving, 234–235
 theory of cognitive ability, 61–62,
 239–244
Accommodation, 380–381
Alzheimer’s disease, 33
Animal magnetism, 258
Aptitude-Treatment Interactions (AIT), 16, 19,
 20
Assimilation, 381–382
- B**
Behaviorism, 205–206
Bio-Ecological theory of intelligence, 65–81;
 see also Intelligence and biology
Blackboard architecture theory, 240, 245; *see*
 also Schematic reasoning
Bodily kinesthetic intelligence, 293
“Bond” theory of intelligence, 121–122
- C**
Capacity, 4–5, 118, 119, 120, 121, 125, 164,
 240–242; *see also* Ability
Cognitive Abilities Measurement (CAM)
 taxonomy, 327–336
 and declarative knowledge, 345–346
 and declarative learning, 341–343
 evaluation of, 330–333
 and information-processing speed, 338–341
 and interaction factors, 351–352
 and procedural knowledge, 347–348
 and procedural learning, 343–345
 and temporal processing, 348
 and time-sharing ability, 349–351
 versus structure-of-intellect model, 329–330
 and working memory, 336–338
Cognitive capabilities, *see also* Ability theory
 of, 151–197
Cognitive components, 308–310
Cognitive components approach, 89, 367
Cognitive correlates approach, 89, 367

Contextualized intelligence, 295–297
 Contextual modifiability, 395–402
 Convergent thinking, 138
 Creativity, 34, 138–142
 and IQ, 139
 and mental illness, 138
 and originality, 139
 and psychoticism, 138–139, 140
 and schizophrenia, 138, 141

D

Declarative knowledge, 345–346
 Declarative learning, 341–343
 Distributed intelligence, 295–296, 301
 Divergent thinking, 138
 Domain, 54, 290, 316, 334
 factors of, 50–61, 261
 in auditory reception, 56
 in cognitive speed, 57–58
 higher order, 58
 in idea production, 56–57
 in knowledge, 58
 in language competence, 54
 in memory and learning, 55
 in reasoning and thinking, 55
 in visual perception, 55–56

E

Elementary cognitive tasks (ECT), 269–279;
 see also Mental speed, Reaction time
 measures
 Essence theory of intelligence, 30, 40, 158–160

F

Factor analysis, 44–47, 90–92, 122–123,
 262–266
 and psychometrics, 69ff.
 Field, 290–291
 Fluid and crystallized intelligence, 21, 46, 47,
 59, 60, 120–121, 163, 165, 172–174,
 196–197, 246–248, 264, 372, 379
 Furneaux's model, 128–131; *see also*
 Information processing, Mental speed

G

The g factor, 65–66, 69–72, 90–92, 211–212,
 266–269, 366–367
 and the brain, 268, 269
 and cognitive ability, 94–111
 General intelligence, 31, 66, 69–72, 210ff., 248

and ability
 models of, 94–112
 and biology, 72–73, 110, 248–252
 declining with age, 184–187, 220; *see also*
 Fluid and crystallized intelligence
 definition of, 90–92
 demographic differences, 228–229
 and development, 175–184
 Essence theory, 158–160
 and factory analysis, 69ff., 90–92, 266–269
 and genetics, 187–195, 215–216, 220–221,
 227
 and heritability, 187–195, 227–228
 importance of, 92–93, 225–229
 and IQ and IQ testing, 70, 93–94, 96,
 108–109
 and psychometrics, 69–72, 266–269
 stability of, 221–223
 theories of, 158–165
 Genetics, 79, 187–195, 215–216, 217–218,
 227; *see also* Heritability
 and the brain, 194–195
 and environmental influence, 188–194,
 217–218
 Gf-Gc theory of intelligence, *see* Fluid and
 crystallized theory of intelligence
 Goal setting, 13, 18–19, 22
 Gravitation, 119, 127
 Guilford's theory of intelligence, 8–9, 98,
 101–108, 121, 168–172, 379–380; *see*
 also Structure of Intellect model
 and structure of intellect model, 168–172,
 329

H

Heritability, 73, 187–195, 227, 228; *see also*
 Ability, Genetics
 and general intelligence, 193–195, 227–228
 and IQ, 193–195
 and twins, 187–188, 189–191, 194
 Hierarchical theory of intelligence, 44–62, 165,
 166–167, 264, 325, 379
 Hume, David, 29–30

I

Information processing, 6, 21, 71, 136–137,
 244, 245, 246, 247, 269–279; *see also*
 Mental speed, Psychometrics, Reaction
 time measures
 automizing, 372–373

- and biological influences, 248–252
 a consensus model of, 310–311
 knowledge-acquisition components, 368–369
 and mental retardation, 88–89
 metacomponents of, 364–366
 and novelty, 370–372
 and performance, 6, 21
 performance components of, 366–369
 speed of, 126–132, 136, 279–282, 338–341,
 372; *see also* Mental speed
- Inspection time measures, 126–127, 132, 133
- Intellectual style
 development of, 390–396
- Intelligence
 as an acquired repertoire, 208–218
 factors of, 210–212
 measurement of, 208–209, 219
 size of, 217, 219
 and attention, 3, 7–10, 17, 20, 22, 186, 187
 as a behavioral trait, 206–230
 and causality, 207–208
 and genetics, 215–216
 measurement of, 214–215
 theory, 218–225
 and biology, 66, 67, 72–73, 117–142, 239,
 248–252, 295; *see also* Bio-
 Ecological theory of intelligence
 genetics, 67–68, 120, 121, 124, 125
 and information processing, 248
 and IQ tests, 67–68
 and brain damage, 288
 and cognition, 233–254
 schematic reasoning, 240–244
 verbal, 244–246
 visual-spatial, 244–246
 theory of, 239–244
 definition of, 118–121, 152, 208, 234,
 258–259
 and the environment, 193, 194, 216–217,
 373–378
 adaptation, 373–375
 selection, 277–278
 shaping, 275–277
 and experience, 369–373
 and evolution, 289
 and learning, 3, 10–16, 55, 125, 289, 295,
 299–302, 339; *see also* Ability, Skill
 acquisition
 declarative, 341–343
 probabilistic, 352
 procedural, 343–345
 implicit, 352–353
 and personality, 29–40, 120, 138–142, 395;
 see also Creativity, Intellectual styles
 and children, 36
 and IQ, 34ff.
 self-ratings of, 35–36
 Intelligence tests, 212–213, 252–253, 268; *see*
 also Test-taking performance
 validity of, 223–225
- J**
- Job performance, 15–16, 18, 213–214, 223,
 298, 395
 and the *g* factor, 266–268
- K**
- Kahneman's theory of intelligence, 7–8
 Knowledge acquisition, 368–369, 381; *see also*
 Skill acquisition
- L**
- Latent traits, 44ff., *see also* Ability
 Linguistic intelligence, 291, 265
 Logical-mathematical intelligence, 292
- M**
- Maximal vs. typical performance, 2, 3–23; *see*
 also Ability-as-capacity
 determinants of, 16–17
 Mental ability, 86–90, 259–260; *see also*
 Ability
 early history of, 86–90
 and factor analysis, 262–266
 and the *g* factor, 266–269
 testing of, 260–266
 Mental models, 236
 Mental representation, 365
 Mental retardation, 88–89
 Mental speed, 80, 89, 126–132, 184–186,
 269–279, 297, 338–341; *see also*
 Information processing, Reaction time
 measures
 and accuracy, 321–322, 340–341
 and decision making, 353
 and general intelligence, 126–127
 Mental test theory, 44, 87; *see also* Ability
 Methodology of testing, 44–48
 factor analysis, 44–47
 Monitoring, *see* Self-regulation

Motivation, 4–5, 15–16, 18, 75–77, 120, 142, 152, 208

Multiple intelligences theory, 287–294
in the school setting, 299–302

Musical intelligence, 291

Myers-Briggs Type Indicator, 37–38

N

Nerve conduction velocity, 272–275

P

Performance, *see* Ability, Maximal vs. typical performance, Job performance

Personal intelligence, 293–294, 298–299

interpersonal, 293, 299

intrapersonal, 293–294

Phenotypic trait, 206

Phlogiston, 257

Polygenetic trait, 193, 215

Positive manifold, 107, 122, 164, 263

Practical Intelligence for School Project, 301

Primary mental ability theory of intelligence, 165, 167–172, 195–196, 379; *see also* Thurstone's theory of intelligence and fluid and crystallized intelligence, 172–173

Procedural knowledge, 347–348

Procedural learning, 343–345

Project Spectrum, 299–300

Psychometrics, 66, 123, 124, 125, 265, 289, 378

models of, 66–68

R

Reaction time measures, 126–127, 132, 133, 269–279, 339–340; *see also* Information processing, Mental speed and the *g* factor, 270–279

S

Schematic reasoning, 240–244

Schizophrenics, 138

Scientific theory, 117–118, 154, 236–277
causal, 238
descriptive, 237

Script-based behavior, 369–370

Self-government, *see* Self-regulation

Self-regulation, 17, 19, 22, 89, 296–297, 366, 384–390
forms of, 386–388
levels of, 388
leanings of, 389–390

scopes of, 389

styles of, 385–386

Sensory discrimination, 125, 126

Situated learning, *see* Contextualized intelligence

Skill acquisition, 11–13, 343–345, 368–369

Spearman's theory of intelligence, 7, 8, 90–92, 95, 99, 122, 159–160, 264, 308, 378–379; *see also* Essence theory, the *g* factor, General intelligence

Stanford-Binet test, 157, 163, 165, 214, 300

Structure of Intellect model, 329; *see also* Guilford's theory of intelligence versus Cognitive Abilities Measurement taxonomy, 329–330

T

Temporal processing, 348–349

Test-taking performance, 3–23, 20–21, 212–215, 223–225; *see also* Ability, Intelligence tests and factor analysis, 44–47, 69 and the *g* factor, 266–268 and genetics, 67–68, 80 maximal vs. typical, 2, 3–23 measurements of, 213–214, 311–327 stability of, 5–6

Thomson's theory of intelligence, 99–100, 122, 164

Thurstone's theory of intelligence, 98, 122, 167ff., 237, 379; *see also* Primary mental ability theory of intelligence

Time-sharing ability, 349–351

Triarchic theory of intelligence, 238, 364–382

V

Verbal intelligence, 244–246

Visual-spatial intelligence, 213, 244–246, 292–293, 348–349, 365

W

Wechsler intelligence scales, 123, 127, 133, 163, 180, 214, 267

Working memory capacity, 12–13, 271, 279–281, 313, 316, 322, 334, 336–338, 342, 344, 346 and mnemonics, 342

Z

Zone of proximal development, 370