DOCUMENT RESUME

ED 142 198 IR 005 050

AUTHOR	Snow, Richard E.
TITLE	Theory and Method for Research on Aptitude Processes: A Prospectus. Technical Report No. 2.
INSTITUTION	Stanford Univ., Calif. School of Education.
SPONS AGENCY	Advanced Research Projects Agency (DOD), Washington, D.C.; Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
PUB DATE	Oct 76
CONTRACT	N00014-75-C-0882
NOTE	88p.
EDRS PRICE	MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS	*Aptitude; Cognitive Ability; Cognitive Processes; Correlation; *Individual Differences; Memory; Research; *Research Methodology; *Theories

ABSTRACT

The theoretical and methodological concepts available to and needed by research on aptitudes as cognitive processes are discussed. Contemporary views of cognitive processes are examined in relation to individual difference constructs and the methods used to examine their reliability and validity. Individual difference constructs are discussed in relation to cognitive process models and research thereon. Studies of short-term memory processes are reviewed to demonstrate the complementary strengths and weaknesses of experimental and correlational methods and concepts. A coordinated approach to the study of aptitude as information processing is suggested. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED142198

THEORY AND METHOD FOR RESEARCH ON APTITUDE PROCESSES: A PROSPECTUS

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

RICHARD E. SNOW

TECHNICAL REPORT #2

APTITUDE RESEARCH PROJECT
SCHOOL OF EDUCATION
STANFORD UNIVERSITY

Sponsored by

Personnel and Training Research Programs
Psychological Sciences Division
Office of Naval Research

and

Advanced Research Projects Agency

under

Contract No. N00014-75-C-0882

The views and conclusions contained in this document are those of the autho's and should not be interpreted as necessarily representing the official policies, either expressed or implied, at the Office of Naval Research, the Advanced Research Projects Office, or the U.S. Government

Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

OCTOBER 1976

TM005050

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 2	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Theory and Method for Research on Aptitude Processes: A Prospectus		5. TYPE OF REPORT & PERIOD COVERED Technical Report #2
		6. PERFORMING ORG. REPORT NUMBER 2
7. AUTHOR(s) Richard E. Snow		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0882
9. PERFORMING ORGANIZATION NAME AND ADDRESS School of Education, Stanford University, Stanford, California		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 154-376
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs, Psychological Sciences Division, ONR, 458		12. REPORT DATE October 1976
		13. NUMBER OF PAGES 65
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

UNLIMITED

18. SUPPLEMENTARY NOTES

This research was jointly sponsored by the Office of Naval Research and the Defense Advanced Research Projects Agency.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Individual differences, cognitive processes, information processing, aptitudes cognitive abilities, experimental and correlational research methodology, short-term memory.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The theoretical and methodological concepts available to and needed by research on aptitudes as cognitive processes are discussed. Contemporary views of cognitive processes are examined in relation to individual difference constructs and the methods used to examine their reliability and validity. Individual difference constructs are discussed in relation to cognitive process models and research thereon. Studies of short-term memory processes are reviewed to demonstrate the complementary strengths and weaknesses of

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

experimental and correlational methods and concepts. A coordinated approach to the study of aptitude as information processing is suggested.

ERIC Full Text Provided by ERIC

Table of Contents

A vast literature in educational psychology attests to the fact that individual differences in learner aptitudes predict learning outcomes, and a substantial body of literature also now demonstrates that aptitude variables often interact with instructional or training treatment variables in these predictions (Cronbach & Snow, in press). These aptitude-treatment interactions (ATI) have important implications for the development of instructional theory and research and for instructional improvement. But if practical and theoretical use is to be made of aptitude information in instructional work, then individual differences in aptitude for learning will need to be understood, at a more analytic level, as individual differences in psychological processes.

An initial report in this series (Snow, 1976) gave a summary of recent instructional ATI studies concerning those aptitude constructs judged most worthy of further research, reviewed two methodological developments relevant to such studies, and then began the task of collating laboratory research on relations between aptitude measures and measures reflecting cognitive processes. It was suggested that a laboratory science of aptitude could be constructed to complement continuing ATI research on instruction by pursuing common process analyses of individual differences in aptitude and learning. The present report reviews the theoretical and methodological foundation on which such a science can build, and sketches some directions in which a process theory of aptitude might be sought.

Orientation

Background

The idea of a laboratory science for the analysis and interrelation of aptitude tests and learning tasks is not new. The problem of individual differences in learning has been of interest periodically in experimental psychology since its early days. Glaser (1967) has reviewed this history. In one form or another new research in this direction has been suggested by several recent writers (Gagné, 1970, Estes, 1970, 1974; Glaser, 1972, 1976). Glaser and Resnick (1972) gave some examples of experiments serving task analytic purposes. Some of the instructional ATI experiments are also useful for such purposes, if reinterpreted as suggesting only possible ATI mechanisms rather than probable generalizations to instruction.

They combine with laboratory studies arising from experimental psychologists' renewed interest in cognitive processes related to intelligence (see, e.g., Resnick, 1976) and from attempts to alter aptitudes through direct training (e.g., Guinaugh, 1971; Jacobs & Vandeventer, 1971ab). These form a loose but growing collection of provocative suggestions. Some use experimental manipulations to examine the construct validity of aptitude measures. Some use aptitudes to examine the construct validity of measures of learning processes, and some generate new conceptions of aptitude and learning as a result. But as yet there has been no systematic compilation of this literature or development of a theoretical framework with which to plan further efforts. The last time an experimental psychology textbook paid extensive attention to individual differences in learning was a 1952 chapter by McGeogh and Irion (although Underwood, 1957, devoted some pages to a discussion of the methodology of experimental research on individual differences). In 1957, Cronbach issued his famous call for unification of correlational and experimental psychology. This was the impetus for the growth of ATI research on instruction through the 1960's. The milestone symposium edited by Gagné (1967) included several views of laboratory research on individual differences in learning, although no substantive connections between that work and ATI research have been attempted. There is still only minor contact between research on aptitude and research on cognition in instruction. (See, for example, Klahr, 1976.)

Periodically over recent years, however, various other writers have proposed one or another general means of combining experimental and correlational psychology (e.g., Owens, 1968; Vale & Vale, 1969; Hunt & Sullivan, 1974). The general implications of person-situation interactions has become an issue of concern in many quarters of psychology. (Compare, for example, Mischel, 1973; McGuire, 1973; Campbell, 1975; Cronbach, 1975). And currently there are new suggestions for the experimental analysis of individual differences and their use in theories of learning and cognition (e.g., Estes, 1974; Hunt & Lansman, 1975; Underwood, 1975). Thus individual differences in aptitude, in learning, and in related cognitive processes seems now to be a topic on the agenda for basic theory and research, both in the U. S. and abroad (see e.g., Flammer, 1975). The time has finally come for combined, concerted efforts.

The present discussion cannot hope to provide a thorough updating of this field, or a thorough examination of all the relevant theoretical and methodological issues. Nor can it consider in detail all domains of aptitude and learning variables. The research project of which this report is a part is conducting a continuing literature review in these areas. Later reports in this series will carry the results of this review.

Starting Assumptions

Definition of aptitude. An aptitude is an individual difference construct, with its associated measures, that bears an hypothesized or demonstrated relation to individual differences in learning in some particular setting. In education, aptitudes are student characteristics that predict response to instruction under a given instructional treatment. In educational research, then, the defining characteristic of aptitude is relation to learning. Measures of "intelligence" or "scholastic ability" identify aptitude because they predict achievement in conventional schooling. Through decades of demonstrations of this prediction, "ability" and "aptitude" came to be thought of as synonyms. But any special ability, cognitive style, personality, motivation, or interest variable that shows relation to learning ought also to be considered as identifying aptitude. There is, then, no traditional domain of differential psychology that should be called "aptitudes" *a priori*. By adopting this broader definition, the field is left open to the study of new and old constructs alike and to hypotheses about combinations of constructs from different traditional domains. More detailed discussion of this definition is given by Snow (1976) and Cronbach and Snow (in press).

Within this broad definition, it is nonetheless true that most research on aptitude for learning has concentrated on cognitive ability, and this report focusses here also. The concept of general mental ability will be a first cornerstone for any theory of aptitude. The central hypothesis of this report and the research program it advocates, is that individual differences in performance on ability tests and learning tasks are manifestations of cognitive processes common to each. Despite historical arguments to the contrary, notably by Woodrow (1946), intelligence is still often defined as the ability to learn. This definition persists because it is parsimonious and intuitively appealing, and because it makes psychological sense. There are theoretical reasons to believe that

ERIC

individual differences in ability and learning derive from the same psychological phenomena whether one takes an environmentalist (e.g., Ferguson, 1954, 1956; J. McV. Hunt, 1961) or hereditarian (e.g., Garrett, 1946; Jensen, 1972) view. And the research often cited in denying the connection is not convincing (Cronbach & Snow, in press). The two disciplines of differential and experimental psychology, focussing on different aspects of the whole, devised different representation systems and terminology for their points of view--one based on static quantities and vectors in mental space, the other on mechanistic functions and group acquisition curves. Progress will now best be served by relegating this division of labor and all that it implies to the historical closet, and by avoiding where possible the limitations of discipline-specific terminology.

With this view as an entry point, several basic propositions about alternative approaches to research on aptitude and its incorporation into theories of learning and cognition can now be added.

Idiosyncracies vs. general laws. Two extreme positions represent opposing and equally counterproductive views about individual differences in aptitude. One holds that there are none of import; learning can be explained sufficiently by general laws applicable to everyone. The other holds that each individual is unique; only idiographic study of the single case can provide understanding. An intermediate position is likely to be most productive. The fact that a variety of individual differences have been successfully measured across persons, and related to a variety of learning outcomes, argues against both extreme positions. General laws can be stated and studied, but included in their study must be an assessment of the boundaries beyond which they cannot be generalized. Idiosyncratic processes can be studied, but included in their study must be some assessment of the possibility of inter-individual similarity. We can expect that both approaches would eventually sort individuals into arrays of relatively homogeneous groups of $1<N<\infty$, one approach by recognizing boundary differences, the other by recognizing similarities.

Typologies vs. multivariate measurement. But sorting individuals into labelled categories according to boundary conditions or similarities identified in one type of experiment breeds an archaic form of thinking about individual differences. Typologies were discarded by modern psychologists when it was recognized that many dimensions were needed to characterize an individual. No type category ever contains individuals homogeneous

ERIC

in all relevant respects. Typological distinctions may define hypothesized dimensions worth further study; there may even be occasions where the bimodal character of an individual difference distribution supports some form of categorical thinking. But multivariate continuous parametric measurement has so far proven to be the most efficient and versatile approach to the problem of studying individual differences of all kinds. Until enough is known about individual differences in cognitive processes to rule out curvilinearity and/or to support hypotheses about discontinuities, typological thinking, and the arbitrary cleavage of continuous variables that it promotes, should be avoided in favor of multidimensional conceptualization and multivariate statistical analysis.

Hypothesis testing vs. estimation of relationships. If individual differences in aptitude and learning variables are viewed as continuous, then it follows that a primary aim is to estimate the form and strength of relationships among them. This view holds as well for the study of relations between experimental variations and individual differences. Significance testing is then of secondary importance--of value as one guide to efficient use of research resources, but hardly the final arbiter of what is substance and what is shadow for the construction of theory.

Complexity. With this position, it must also be noted that one-to-one correspondences between present aptitude and learning constructs will not likely be found. What is to be sought is some kind of mapping of each set of constructs onto the other. Relationships in this mapping are unlikely to be simple.

Causality. Particularly to be avoided in such a mapping are assumptions that place aptitude constructs at either more or less "basic" a level of understanding than learning or other process constructs. Individual difference measures, whether based on cognitive test scores or laboratory task parameters, do not automatically reflect "fundamental unities." One kind of measure cannot be routinely taken as providing causal explanation of the other. While it is possible, for example, that some aptitude constructs may come to be explained as complex functions of "more basic" information processing constructs, it is also possible that other aptitude constructs will be found to reflect rather directly some "more basic" biological features of the individual which in turn control information processing.

Generalization. Finally, conceptions of aptitude in learning built from laboratory analyses should not be expected to generalize directly to instructional settings. Laboratory models are analogs which can enrich conceptualizations of aptitude in instruction and can suggest improvements in aptitude measurement there. Ultimately, however, a theory of aptitude will need to be a theory of aptitude *in situ*.

Existing Models

A means of thinking about individual differences is required that promotes hypotheses about the cognitive processes that may be involved in and may help distinguish aptitude constructs. It seems reasonable to look first at existing theoretical formulations. The models used in the factor analytic tradition, in S-R association theory, and in information processing conceptions of cognition, may offer useful starting points even if none of these constructions deals explicitly with process conceptions of individual differences. Some theoretical or methodological coordination of parts of these models may still be possible.

Factor Analytic Models

While factor analysts may have been the first modern cognitive psychologists (Carroll, 1976), the differential psychological tradition has not moved much closer to theories of cognitive processes than the production of taxonomies for classifying measured individual differences. Guilford's (1967) Structure of Intellect model is called "informational", but it does not give an account of information processing. Hierarchical views of intelligence (e.g., Vernon, 1965; Cattell, 1971) help assure parsimonious classification and interpretation of abilities, but also fall short in providing process models. Investigators of cognitive "styles" (e.g., D. Hunt, 1975; Kagan & Kogan, 1970; Witkin, 1973) have aimed at process theories, but style constructs have typically been studied in isolation, disconnected both from the rest of differential psychology's catalog and from experimental cognitive theories.

Hierarchical organization of abilities. Even if it does not yield process concepts, the hierarchical model of ability organization seems to be the most reasonable starting place for research in this area, from a differential psychological view. Figure 1 shows a structure designed to approximate those fashioned by Vernon (1965), Cattell (1971), and Cronbach (1970), and to be consistent with Guttman's (1965) multidimensional

scaling of Thurstone's data as well as other factor analytic research. A review by Horn (1976) of recent factor analytic studies also supports the major distinctions shown and their presumed levels in the hierarchy. A general mental ability (G) is displayed at the top, with a division into fluid analytic ability (G_f) on one side and crystallized verbal ability and educational achievements (G_c) on the other. Horn's review suggests that spatial visualization abilities have been found broad and coherent enough to be listed at the general level (G_v) and that verbal productive thinking and fluency abilities can be given a place not far below G_c.

- - - - - - - - - - - - - -

Insert Figure 1 about here

- - - - - - - - - - - - - -

Below these, it is possible to sketch in more specific abilities and skills as subsidiaries to one or another of the more general abilities. Such a structure is schematic; there need be no firm commitment to exact details. There is more evidence supporting some constructs than others, however, and this is suggested by the solid and dashed lines, respectively. In general, the crystallized side of the hierarchy seems to have been more studied and is thus more fully elaborated. This might suggest that one major focus of future research should be on the G_f and G_v regions of the hierarchy.

Guilford's Structure of Intellect (1967) is a facet model, wherein each ability is defined as a combination of one of five kinds of mental operations (cognition, memory, convergent thinking, divergent thinking, and evaluation), on one of four kinds of content (figural, symbolic, semantic, and behavioral), with one of six kinds of products (units, classes, relations, systems, transformations, and implications). As such it is superficially inconsistent with the hierarchical view. But some recent reanalyses (see Cronbach & Snow, in press; Merrifield, 1970) suggest that Guilford data can fit a hierarchical model. It is thought, for example, that Guilford's product categories are distinguishable within categories defined by sections of his operation × content matrix. There is also the hypothesis, though not substantiated, that the product dimension itself should be thought of as hierarchically ordered. Thus Guilford's model, which posits more than 120 distinct special abilities, might with further work be made to fit into Figure 1. (See, e.g., Haynes, 1970, but also Humphreys, 1962.)

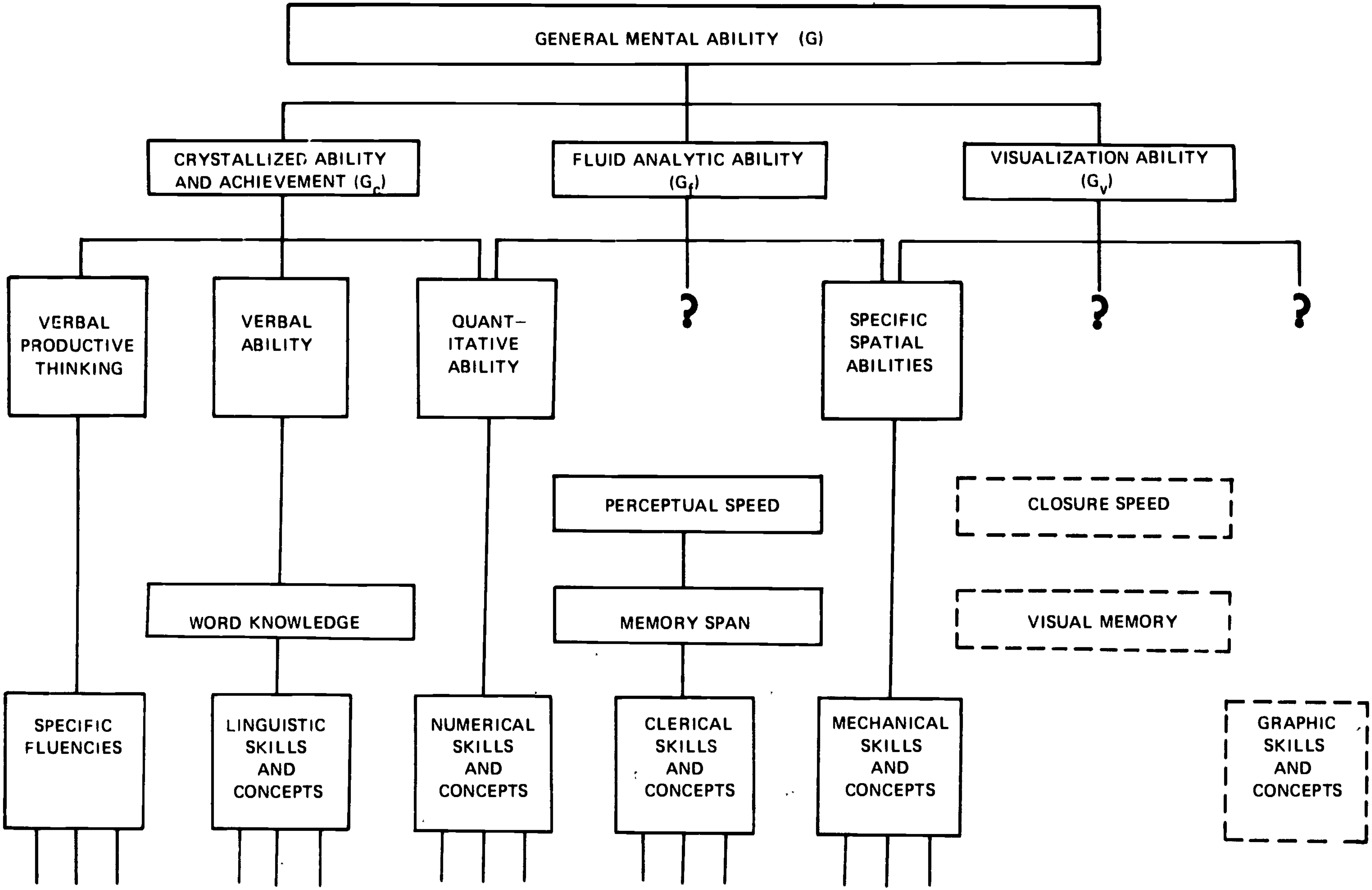


FIG 1. A HIERARCHY OF MENTAL ABILITIES

Parsimony. As implied above, a principal reason for adopting some such hierarchy as a starting point is one of parsimony. We wish to deal in theory with as few aptitude constructs as possible and should allow additional constructs only as data dictate. By starting with the most general constructs, continuing experiments can then show whether and when distinctions between subabilities are required. The kind of hierarchy shown in Figure 1 has been elaborated through decades of factor analytic research to the point where, to be complete, about 100 ability factors would have to be located somewhere in the body of the figure. Rather than accepting this degree of differentiation as given, we should prefer to determine if an alternative method of research on aptitude would yield a similar set of distinctions. Presumably, laboratory experiments can be designed to help build and test process models of such distinctions. The hierarchy will help organize that research and its results in the most parsimonious way.

As research continues, we will be faced with correlations between, say, a spatial, or numerical, or verbal ability measure and some learning measure. In the hierarchical view, one cannot interpret the relation in terms of the special ability unless it can be shown that a more general construct does not account for the result. The argument applies with particular force to the ability measures used in research on cognitive processes, but should apply as well to process measures themselves. The goal ultimately is to reach the minimum set of the most general constructs needed to account for individual differences in cognitive processes in learning. This proposition, of course, is meant to limit interpretation, not hypothesis generation.

Other properties of the hierarchical factor model. The hierarchical factor model carries several other useful properties, in addition to providing a parsimonious schema. Any factor analysis of task performances classifies together those tasks that seem to correlate highly relative to their correlations with tasks classified into different categories. Factors, then, are at least classification principles (Vernon, 1951). Great theoretical weight should not be placed on the classification of tasks produced by any one factor analysis. When similar classifications persist over decades of research, however, one is justified in investing research effort in seeking a deeper explanation for the persistent task relations that produce them. The hierarchical model tends to separate those classifications that consistently appear, at the more general level, from the

narrower, more specialized classifications that may flicker in and out. The large factors, and the more persistent of the specialized factors, provide a network of reference constructs summarizing a great many task interrelationships. Any new experiment seeking to analyze particular cognitive processes benefits in generalizability and external validity by bringing in this network of task relationships for background reference purposes. The new cognitive task of primary interest can be examined in the context of its correlates in the existing network. The hierarchical model allows us, in addition, to choose the level of ability specialization we wish to represent in a study.

Further, to represent the network of reference abilities in any new study, factor composite scores will usually be preferable to scores derived from single tasks, because the former will typically be more stable and reliable as measures than will the latter. If constructed properly, they provide a succinct mathematical summary of between-task relations for each ability construct and an effective means of statistical control in reasoning about partial correlations.

Still another useful aspect of the reference constructs represented by tests is their empirical connection to real world performances. A new laboratory task has no empirical connection to the world outside, although it may have theoretical relevance. Since many of the tests used to mark reference constructs have long histories of relation to practical criteria, the network provides tracks through which hypotheses relating laboratory and field phenomena can be traced.

Hypotheses. The vertical dimension of the hierarchical model represents differences in referent generality among ability constructs (Coan, 1964). Constructs at higher levels typically refer to larger classifications of tasks, with more real world connections, that are likely to transfer to, correlate with, or otherwise generalize to, a broader range of other performance measures. Traditionally, they have been thought to reflect deeper, more fundamental properties of intellect as well. And the general factors account for most of the ATI findings of previous instructional research (Cronbach & Snow, in press).

Even the large persistently-appearing reference constructs need not reflect fundamental cognitive phenomena, of course. Tests, like all tasks, can correlate for superficial as well as fundamental reasons.

ERIC

Many traditional factor analysts prematurely concluded that their factors reflected basic psychological units of cognition, and turned inward to concentrate on the refinement of their methodology. The contrast between factors interpreted as fundamental unities vs. classification principles is nicely captured in Cronbach's (1970, p. 330) astronomical analogy. Does factor analysis discover new "planets" in ability space by analyzing the behavior of already discovered planets and their interrelations? Or does it identify "constellations" which are merely convenient clusters we can use in mapping that space? Reference constructs in the hierarchical model perform the latter service at least, while not ruling out the possibility that key parts of some constellation may indeed represent fundaments deserving closer inspection. Table 1 gives a listing of the definitions usually offered for each of the major abilities identified in Figure 1. These may serve as guides to the development of more detailed process hypotheses, and help to maintain an historical perspective as research proceeds.

- - - - - - - - - - -

Table 1 about here

- - - - - - - - - - -

In addition, the vertical general-to-specific continuum may itself reflect a fundamental feature of the structure of cognition. Guttman's (1965) rescaling of Thurstone data suggests a certain centrality for the more "analytic" abstract ability tests. These are not only more abstract tasks than those found in the periphery of the scaling and labelled by him "achievements"; they are also more complex as information processing tasks. This complexity dimension in the hierarchy deserves research attention regardless of the labels used to identify particular factors or categories of tests.

Thus the factor analytic tradition provides a network of reference ability constructs with which we can examine the external validity of laboratory analyses, and some formative hypotheses that may deserve such analysis.

TABLE 1

A Compilation of Definitions of Major Human Mental Abilities

General Mental Ability (G)

Binet (in Terman, 1916, p. 45)	"The tendency to take and maintain a definite direction; the capacity to make adaptations for the purpose of attaining a desired end; and the power of auto-criticism."
Binet & Simon (1916, pp. 42-43)	"... judgement, otherwise called good sense, practical sense, initiative, the faculty of adapting one's self to circumstances. To judge well, to comprehend well, to reason well, these are the essential activities of intelligence."
Spearman (1923)	"... everything intellectual can be reduced to some special case ... of educing either relations or correlates." (p. 300) Eduction of relations--"The mentally presenting of any two or more characters ... tends to evoke immediately a knowing of relation between them." (p. 63) Eduction of correlates--"The presenting of any character together with any relation tends to evoke immediately a knowing of the correlative character." (p. 91)
Stoddard (1943, p. 4)	"... the ability to undertake activities that are characterized by (1) difficulty, (2) complexity, (3) abstractness, (4) economy, (5) adaptiveness to a goal, (6) social value, and (7) the emergence of originals, and to maintain such activities under conditions that demand a concentration of energy and a resistence to emotional forces."
Freeman (1955, pp. 60-61)	"... adjustment or adaptation of the individual to his total environment, or to limited aspects thereof. ... the capacity to reorganize one's behavior patterns so as to act more effectively and more appropriately in novel situations. "... the ability to learn. ... the extent to which [a person] is educable. "... the ability to carry on abstract thinking ... the effective use of concepts and symbols in dealing with ... a problem to be solved."
J. McV. Hunt (1961, p. 362)	"... conceived as intellectual capacities based on central processes hierarchically arranged within the intrinsic portions of the cerebrum. These central processes are approximately analogous to the strategies for information processing and action with which electronic computers are programmed."
Jensen (1969, p. 9)	"When the term 'intelligence' is used it should refer to [Spearman's] g , the factor common to all tests of complex problem-solving"
(1970, pp. 147-148)	"... mental tests can be ordered along a continuum going from simple to complex. ... The intercorrelations among tests are roughly related to their degree of proximity on the complexity continuum, and tests which are intended to identify g ... show increasing correlations with other tasks as one moves along the continuum from simple to complex."

ERIC

TABLE 1 (Continued)

	Crystallized Ability (G_c)
Cattell (1963, p. 2)	"Crystallized ability loads more highly those cognitive performances in which skilled judgement habits have become crystallized (whence it's name) as the result of earlier learning application of some prior, more fundamental ability in these fields. Thurstone's Verbal and Numerical primaries, or achievement in geography or history, would be examples of such products."
Horn (1976, p. 445)	"Awareness of concepts and terms pertaining to a broad variety of topics, as measured in general information and vocabulary tests and in tests which measure knowledge in science, mechanics, social studies, English literature, mathematics, and a variety of other areas. It is also manifested in the Information, Vocabulary, Comprehension, Similarities and, to a lesser extent, Arithmetic subtests of the Wechsler Scales ... . In much of the British' work it is labeled verbal-educational (v:ed) intelligence."
	Fluid-Analytic Ability (G_f)
Cattell (1963, p. 3)	"Fluid general ability, on the other hand, shows more in tests requiring adaptation to new situations, where crystallized skills are of no particular advantage."
Horn (1976, p. 445)	"Facility in reasoning, particularly in figural and non-word symbolic materials, as indicated in tests such as letter series, matrices, mazes, figure classifications, and word groupings, as well as the block designs, picture arrangements, object assembly, and picture completion subtests of the Wechsler Scales... . Some characterize it as non-verbal intelligence (although verbal tests can measure it) or performance IQ. In the British work it is known as spatial-perceptual-practical intelligence (k:m)."
	Visualization Ability (G_v)
Horn (1976, p. 448)	"When analyses pertain to concepts more general than the primary abilities, the various spatial tasks ... [involving ability to perceive and transform images of spatial patterns, maintaining orientation in spatial arrangements] tend to hang together in what can be referred to as a general visualization dimension which seems to be at least somewhat distinct from G_f, and is clearly distinct from G_c ... ."

NOTE: Definitions of various special abilities at lower levels of the hierarchy shown in Figure 1 may be found in French, Ekstrom, and Price, (1963).

Association Models

A better way to arrive at a process theory of aptitude may be to start with a basic theoretical framework derived from experimental psychology (Melton 1967). This might suggest some process hypotheses about individual differences directly. It would at least offer an orderly way to use individual difference data in testing and elaborating the existing experimental framework (Underwood, 1975). And obtained relations between aptitude measures and cognitive process measures could suggest a coordinated model.

An initial framework. Early associationist views interpreted intelligence as simply the sum total of S-R connections brought forth through an individual's learning history, no more structured than E. L. Thorndike's "pile of leaves". Modern theorists, however, separate internal phases of the S-R connection and infer hierarchies of increasingly complex associative networks. Melton (1967) distinguished stimulus differentiation and integration processes that transformed the stimulus (S) into a stimulus-as-coded (s_1), the mediating responses (r_m) that hooked s_1 to R, and then a class of response discrimination and integration activities. It was suggested that individual differences in learning be sought within such a categorization. His diagram implies further that mediation is an alternative hookup between S and R , such that qualitatively different connections might occur in different individuals.

Hierarchical coordination. Hierarchical models have been built on an associative base, from Hull (1934) through Spence (1959) to recent, more elaborated views. Gagné's (1970) hierarchy of types of learning provided both a general classification and a means of analyzing instructional tasks into hierarchical arrangements of prerequisite steps. In his analyses, relations to ability constructs typically were located at the base of such structures, although he has also drawn hierarchical analyses of some ability constructs themselves (e.g., Gagné, 1968). Berlyne (1965) hypothesized internal chains of situational and transformational responses, linking the external S and R, and arranged in habit family hierarchies. The transformational habits were especially important in Berlyne's thinking. Referring to Spearman's (1923) concept of intelligence, he likened the "eduction of relations" to transformation-selecting habits and the "eduction of correlates" to transformation-applying habits.

Combining these views, it is possible to obtain at least a gross coordination between aspects of the associationist and factor models. Figure 2 depicts one such coordination, of some of Melton's, Gagné's, and Berlyne's ideas with the hierarchical view of parts of Guilford's model mentioned previously. Starting with Melton's basic diagram, each higher level of the figure shows a step up Gagné's hierarchy, represented as an additional link in Berlyne's transformational chain, and associated with a higher ability product in Guilford's system. One might think of separate constructions of this sort for each of Guilford's "content" categories. Some parts of the "operation" facet, however, can be mapped into any one of these figures. Three such categories are shown at the top of the figure, to demarcate three hypothetical phases in the structure: cognition (i.e., discovery or rediscovery, mainly of units, classes and relations), production (i.e., convergent and divergent productive thinking involved mainly in patterns, systems, and transformations), and evaluation (of implications as well as of the adequacy of any other response product).

Insert Figure 2 about here

A schema of this sort is gross and oversimplified, perhaps even far fetched. Certainly the steps should be thought of as categories of cognitive events rather than particular responses. And correlational work has not so far justified many of the ability distinctions implied (Cronbach & Snow, in press). Nonetheless, it leads to some potentially useful hypotheses.

Hypotheses. First, because Guilford identifies only six product levels compared with Gagné's eight levels of learning, two additional kinds of individual differences in ability products can be hypothesized (in parentheses). One of these, the orienting reflex associated with signal learning, has already been studied extensively (Maltzman, 1967). The other, called here "patterns", has not been directly investigated. It is located at the level Gagné called "multiple chaining", where additional stimuli presumably require coordinated patterns of multiple discriminations. Further correlational research might seek to distinguish these two kinds of individual differences in the product sequence.

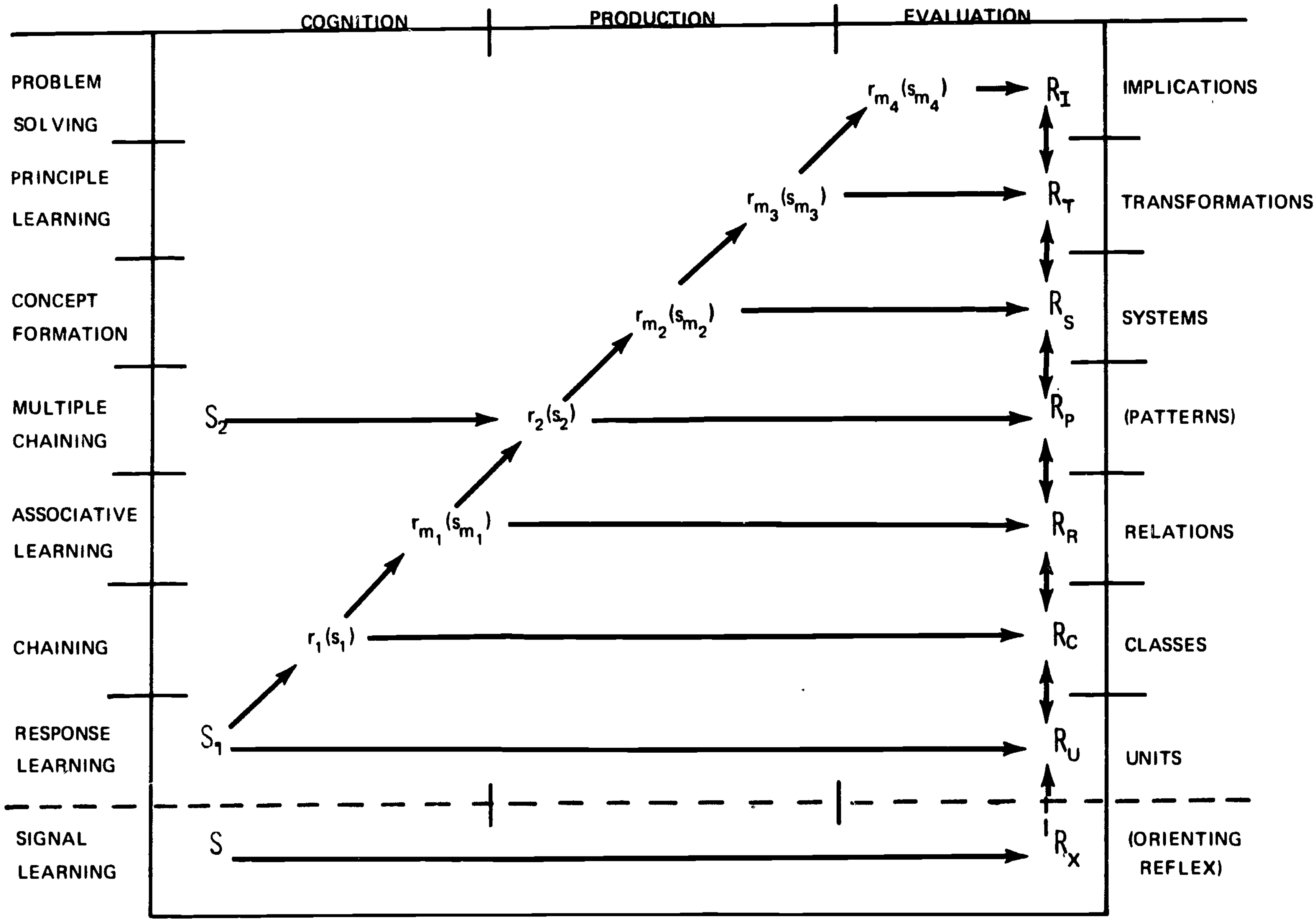


FIG. 2. SCHEMATIC COORDINATION OF LEARNING AND ABILITY HIERARCHIES

The ordering of steps implies further that abilities representing these operation × product cells from the Guilford model both reflect prior learning up to the associated Gagné level and predict individual differences in new learning at those levels. One could imagine the design of learning tasks to represent several adjacent levels or stages of learning and their use in research aimed at correlating individual differences in learning at each level with ability tests chosen from the corresponding response product hierarchy. The hypothesis would be that measures of general ability would correlate with learning in all stages, but particularly in later more complex stages, if the task began at the simple response learning level. Specific tests chosen to represent particular products would correlate primarily with performance at the associated learning level. (See Cronbach & Snow, in press, Chapter 5, for a discussion of the few prior studies on this kind of hypothesis. Note also that this hypothesis can be viewed as a complexification of Jensen's, 1969, concept of Level I and Level II learning abilities. Here, however, there is no specification regarding genetic origins.)

Figure 2 also implies a continuum of increasing complexity of cognitive events (or perhaps, of depth of processing) in learning and ability, running diagonally through the 3 × 8 lattice based on Guilford's categories of abilities. The simplest ability tests and learning tasks then would be those involving cognition of units, classes, and relations; more complex tests and tasks would involve production of systems and transformations, with evaluation of implications being the most complex. Correlations among measures representing each of these cells might be expected to show a structure similar to Guttman's simplex, perhaps with some distortion due to evaluation of response adequacy being involved in all product levels. One might even speculate that a general factor extracted from a matrix of intercorrelations among such tests would correspond to this complexity continuum (following Jensen's 1970 hypothesis).

Still another, related, hypothesis is that the 3 × 8 Guilford matrix can be collapsed into two areas: cognition of lower order products and production of higher order products. It is interesting to note that most verbal tests of crystallized intelligence (G_c) would be classified as cognition of units, classes, or relations, while most figural tests of fluid intelligence (G_f) would be classified as systems, transformations,

or implications. The tests most often taken as measures of a general intelligence factor--Verbal Analogies, Thurstone Letter Series, Raven Matrices--are almost always tests found in the middle categories of relations, systems, and perhaps patterns.

These are mere speculations that might be explored in future correlational studies using factor analytic and/or multidimensional scaling techniques. Beyond these gross connections there has been little interlacing of concepts from associationist and factor analytic models. Association theorists typically viewed individual differences as affecting only the constants, not the form, of general laws, so reference to individual differences in the theoretical writing of the associationist era has been rare. There are, however, some methodological points available from more recent thinking that deserve mention here. They suggest how parts of Figure 2 or other such schemes might be examined experimentally.

An S-R-R paradigm. Glanzer (1967) renewed the consideration of what Spence (1944) earlier called R-R theory, stating the general requirements that any psychological theory, whether S-R or R-R, should specify a process, within an individual, that is accessible by experimental manipulation. This leads to a form of theory called S-R-R (Schoenfeld & Cummings, 1963), where the middle R is a measure of some hypothesized intervening covert response r . Correlational studies are of the R-R type, but they are uninteresting in Glanzer's view because they do not meet the stated requirements for theory. An S-R-R paradigm, however, might yield theory combining correlational and experimental constructs.

Glanzer summarized his experiments on a verbal loop hypothesis in perceptual reproduction tasks to exemplify this form of research. If subjects translate figures into covert verbalizations for storage until response time, then complex or poorly organized figures will require longer verbal codes, and will be less fully encoded under conditions of short exposure. Accuracy of reproduction for each of a set of figures after brief exposure was measured with one group of subjects, and length of verbal descriptions of each figure was obtained using another group of subjects. The hypothesis and one result are diagrammed in Figure 3a. R_1 is a measure implied by the hypothesized r, and the obtained R_1-R relation is consistent with the hypothesis. With this relation in hand, a second experiment (this time on reproducing strings of numbers) showed that three kinds of encoding training (ad lib verbal coding, specified

ERIC

verbal coding, and numerical coding) produced consistent differences in accuracy (although not in the predicted direction).

Insert Figure 3 about here

A conventional experimental approach of this sort is not satisfactory for our purposes. Glanzer's hypothesis does indeed specify a process, in the individual, that can be experimentally manipulated. But the methodology of the example studies does not preserve this specification. R_1 was measured as an average description length for each figure in one group of subjects, and the R_1-R relation was obtained by correlating this average measure with average accuracy obtained in another group of subjects; N was the sample of figures. Thus the result concerns average differences among figures, not individual differences among human subjects.

For research on individual differences in psychological processes, an additional requirement must be added to the S-R-R paradigm: the two R variables must be measures applied to the same subjects, and yielding individual scores for each subject. Then, effects of experimental manipulations of S are reflected in changes in the R-R interrelations as well as in the mean of each R. The middle R might be truly an intermediate response measure in an experiment, or it might be a measure of an individual attribute taken before the occurrence of S. All experiments on ATI are instances of the latter type: the aptitude, while measured before treatment, is presumed to represent individual differences on an intervening variable essential for learning. (Some might prefer to label this latter, ATI case, an instance of R-S-R theory.)

A study by Gavurin (1969) serves as a good, simple example of this form of research. He asked subjects in one condition to solve anagram problems, with letters presented on tiles movable about the table top; in another condition the tiles were taped together, forcing subjects to find solutions mentally. A measure of spatial ability (the Paper Form Board test) administered prior to the experimental manipulation correlated .54 with anagram performance in the latter condition but -.18 in the former condition. The hypothesis and results are diagrammed in Figure 3b. Presumably, the spatial ability test taps individual differences in a process required in purely mental solution of anagrams, but not required

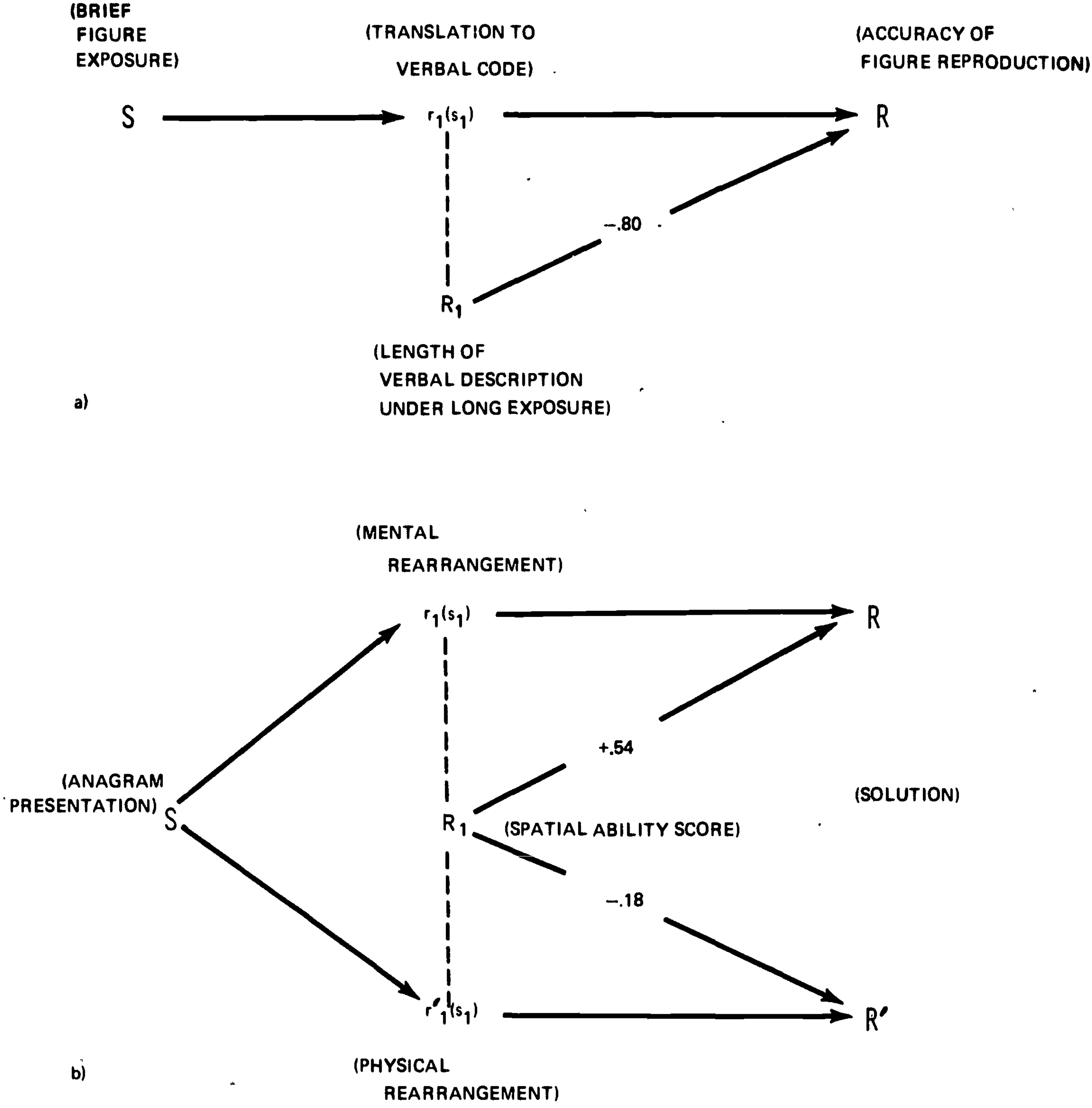


FIG. 3. SCHEMATIC REPRESENTATION OF a) GLANZER (1967) DATA AND b) GAVURIN (1969) DATA IN AN S–R–R PARADIGM.
NUMBERS SHOWN ARE CORRELATION COEFFICIENTS.

when subjects can try out hypotheses externally by manipulation of the stimulus letters on the table-top. Here, all of the above requirements for S-R-R theory are met. Many other experiments reviewed by Snow (1976) fit the same form. Some others, notably Frederiksen (1969), include both ability tests prior to the experiment and introspective strategy measures following it, to represent intervening R variables.

Using such a paradigm, experiments could be constructed to contrast, for example, performance under instructions to rehearse paired associates vs. instructions to imagine meaningful connections between the pairs (following Rohwer's, 1975, work). Individual differences in the latter condition might be expected to relate to ability tests measuring pattern or concept production, while performance in the former condition should correlate mainly with ability tests representing lower product levels. As we shall see, the S-R-R paradigm is not limited to associationist thinking. It provides a methodology applicable in cognitive psychology generally.

Information Processing Models

Modern cognitive psychology is now dominated by an information processing approach based largely on computer models. (See, e.g., Reitman, 1965; Newell & Simon, 1972.) Estes (1975) has traced the recent, rapid development of this approach and its importance for the psychology of human learning and cognition. Anderson and Bower (1974) provided a particularly convincing argument for the relative power of such models over simpler associationistic ideas in accounting for the apparent character of complex learning and related cognitive functions. They note that the information processing approach is essentially a methodology for theorizing about the complex of cognitive machinery connecting stimulus and response, rather than simply a methodology for experimenting on these connections. As such, it allows us to build a theory at once more concrete and more abstract than the rather raw empirical constructs of either the associationistic or factor analytic traditions. Most important for the present purpose, it offers, potentially, a richer basis for conceptualizing the relations among constructs from these domains.

An initial framework. Bower's (1975) overview provides an outline of the major components of most information processing models. This is a suitable framework within which to consider individual differences in

information processing. Bower's listing of possible process components begins by distinguishing the initial perceptual system, short term memory (STM), intermediate or working memory (ITM), and long term memory (LTM). Within these there are subroutines. Sensory buffers register stimulus patterns for brief periods while feature detectors operate to analyze incoming information, leading to pattern recognition. The stream of stimuli is segmented and a scene analysis is built up through successive fixations. STM preserves the temporal and spatial order of the stimuli, and recognizes groupings or chunks, as they are coded or recoded for memory or use in achieving some goal. Rehearsal takes place here, and there are presumed to be differential rehearsal and deletion strategies that are brought to bear, depending on task, instruction, and individual differences. While STM allows active, fast access to incoming information by other central processing components, it has limited storage capacity and duration. A second memory system, ITM, maintains information about the local setting or task environment in which active STM is focused. It updates a "model" of the situation as new information is processed through STM. LTM is the repository of permanent propositional and conceptual knowledge, cognitive and motor skills, attitudes and beliefs, etc. Information is presumed to be organized and reorganized into networks for storage and retrieval here.

Hierarchical coordination. The information processing system as a whole is conceived as a hierarchical organization of programs and subroutines. There is an executive program that monitors subordinate processing stages, choosing, maintaining, and changing programs according to goals. This hierarchical view has in turn led to the hypothesis that tasks differ in the "depth" or complexity of processing they require of the system. Often, the components at a given level in the hierarchy are depicted as arranged in a flow chart to show the sequence of hypothesized steps or stages required in cognitive task performance. Each box in the chart can be elaborated into its own complex flowchart at the next lower level of hierarchy. The programs in this hierarchy are usually described as plans or as production systems (see Newell & Simon, 1972).

As a speculative exercise, we can translate the implications of Figure 2 and discussion of it into information processing terms, adding the features identified by Bower, to produce Figure 4. Included here, then,

are STM, ITM, AND LTM, and the implication that the levels of Figure 2 are cognitions or productions derived by addressing LTM directly or constructed from new work in STM. The evaluation system is equated here with the executive or program monitor, and is most closely associated with ITM, the current conception of the environment. Because most pathways in such a chart would have to be two-way streets, we can do away with most of the arrows usually used to show information flow.

- - - - - - - - - - - - - -

Insert Figure 4 about here

- - - - - - - - - - - - - -

Figure 4 suggests, again, that major distinctions among ability factors can be mapped into the major distinctions of an information processing conception. While pleasing, at least to some eyes, however, constructions such as Figure 4 remain gross and static conceptions. One still needs to analyze more specifically any complex learning, ability, or problem-solving task into the sequences of operations and memory functions presumed to constitute performance on that task.

Hypotheses. Carroll (1976) moved in this direction by performing a task analysis of 24 ability tests. His analysis reaches an intermediate level of hypotheses connecting abilities and combinations of information processing operations. And many hypotheses are generated on which to base further experimental and correlational analysis. But specific measures of individual differences in information processing must still be derived.

A list of all the kinds of individual differences that might exist in such systems would be very long indeed. And many of these hypothesized variables might not be operationalizable at the present time, or at all. We need some means of identifying and attacking the most important of these individual differences, while still keeping at least the major categories of all the possible differences in view.

One possible starting point is a set of elementary information processes such as that proposed by Newell and Simon (1972). The set would consist of the minimum processes required to produce any symbol-manipulating or problem-solving computer system. If these processes were truly elementary, all other observable individual differences in system functioning might be understood as combinations of these. The Newell-Simon

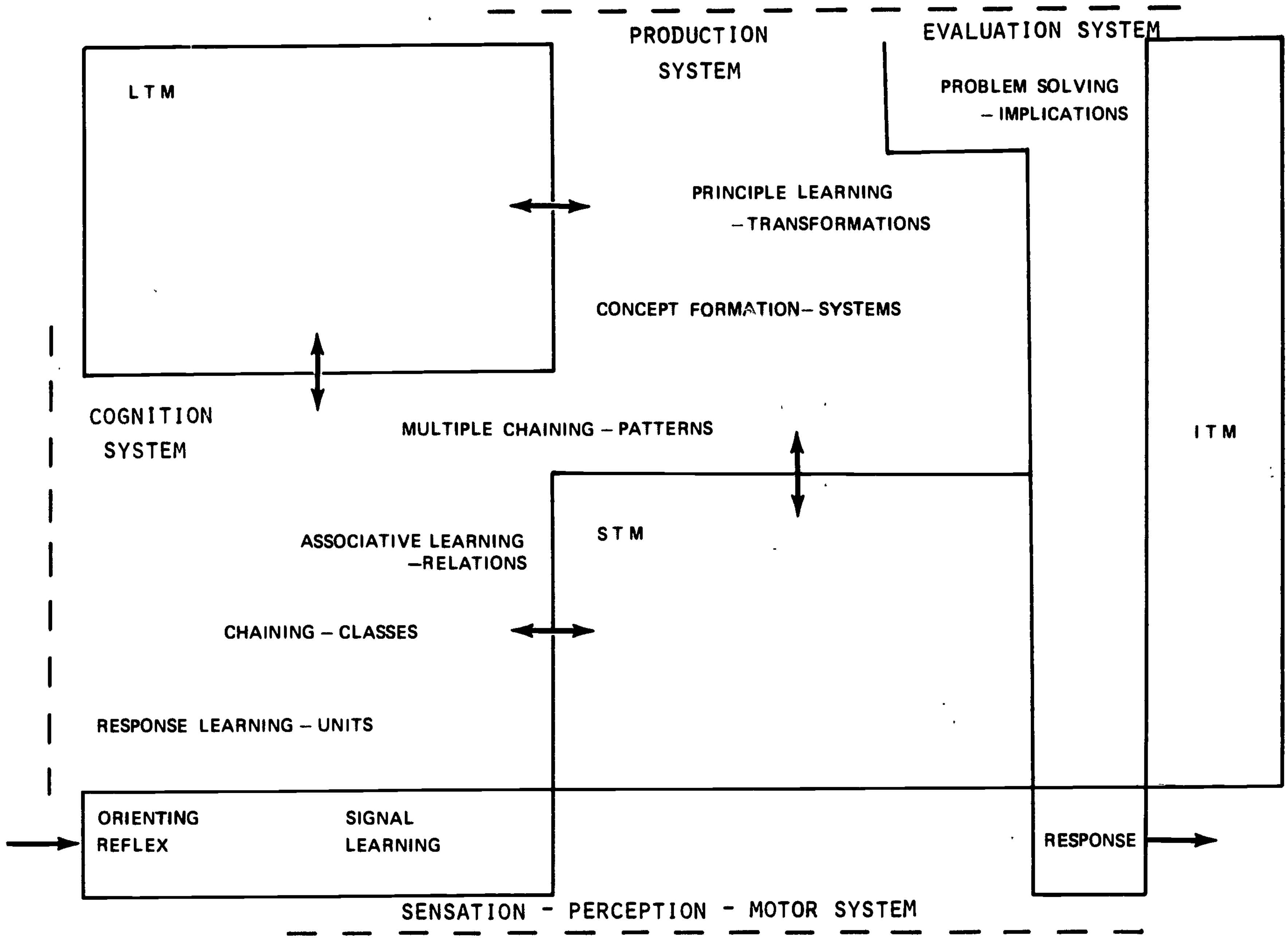


FIG. 4. SCHEMATIC INFORMATION PROCESSING SYSTEM

list includes such elements as discrimination (the capability to alter behavior depending on symbols in STM and to transfer control among alternative programs), comparison (the capability to determine that two symbols are or are not identical or of the same type), symbol creation, reading and writing externally, etc. One could imagine measures designed to assess individual differences in each of these elements independently.

But there seem to be two problems with such an approach. First, these elementary operations are probably accomplished in many different ways or different combinations, both in different systems and also in the same system as a function of task variations. The reduction of task performance to such units appears almost as formidable a problem as does reduction to physiological units. Second, the list of elements concerns central processing; it assumes uniform, high grade input from sensory buffers and from LTM. However, individual differences in "comparison ability", for example, might depend on individual differences in visual acuity and knowledge of Chinese, if the symbols being compared were Chinese characters. With simple Arabic numerals, on the other hand, individual differences in this capability might arise more or less from pure speed of stimulus matching, depending on exposure time. And a given individual might match Chinese characters through painstaking analysis and feature comparison, using efficient visual templates or verbal codes for arabic numerals. Thus, a small basic list of standard measures is unlikely to capture many of the important aspects of individual differences in complex processing.

Furthermore, the information processing approach has concentrated to date on building mathematical or computer models for particular cognitive tasks. The models and the measures derived from them are thus quite task specific. The computer simulations of complex, "slow" processing tasks, while identifying some apparently common structures across tasks, do not compare easily at more detailed elemental levels. Nor do the mathematical models of "fast" processing tasks. There are then many highly specialized models, one for each of the fifty or more kinds of tasks analyzed so far. Generalization of constructs across tasks remains difficult, if not impossible.

Some information processing models do provide an empirical means of generalizing. Each such model typically specifies parameters to represent the speed, efficiency, or capacity of operation on particular processing steps in its task. For example, one set of such parameters, defined by Atkinson, Brelsford, and Shiffrin (1967) for their continuous paired-associate task, are the following: α , the probability of entry of an item into STM; r , the number of items that can be held in STM at one time; θ , the rate of transfer of information into ITM; and τ , the rate at which information becomes unavailable from ITM. Another example would be Sternberg's use of slope and intercept parameters, derived from reaction time data in a memory search experiment. The slope indicates reaction time changes as a function of the number of items in a set stored in STM and is interpreted as the time required to access a single item from the set and compare it with one just presented. The intercept shows performance with a one-item memory set, indicating individual differences in basal reaction time, including time for stimulus encoding, response production, etc. A variety of other kinds of parameters are possible. Any given parameter could be located in a general flow chart depending on what processing step or location it presumably represents. Within each cell, at each step, we might expect measurable individual differences in such parameters to reflect fundamental operating characteristics of human learners. Then, specific parameters might be related across tasks to test their generality. Such an approach would rely on traditional correlational methods of analysis to identify constellations of task parameters that consistently correlate, and to relate these to tested abilities.

The identification of such constellations of parameters and test scores would be helpful, just as the constellations of tests yielded by factor analysis are helpful. But such analyses, alone, are unlikely to yield a new form of theory of aptitude.

Further, the simple task models that yield the parameter measures rest on several assumptions which, if compromised, would lead us to expect individual differences arising from several sources in addition to parameter variations on particular steps. One assumption is that cognitive processes are organized in a train of temporally-ordered independent stages such that each stage begins only when the preceding stage is completed. Another is that all individuals pursue the same

sequence of stages. In this view, one studies mental operations by comparing tasks that differ only in that one contains an additional inserted operation beyond those contained in the other. It is assumed that changes in a task insert or delete processing stages without altering other stages or, at least, that such changes influence the duration of a certain stage without altering others. The effects of task factors that affect independent stages are thus presumed to be additive. These are old ideas in psychology, given new force by the work of Sternberg (1969). Others have elaborated upon the basic assumptions. J. Reitman (1972), for example, proposed that each successive stage, being of a higher cognitive level, takes more time. Thus, in front of each processing step, a queue forms composed of items of information waiting to be processed through that step. Information is lost from these queues due to memory decay and to the need to drop overloads.

But such assumptions may or may not hold for each individual on a given task (Calfee, 1976). Task factors and subject factors may interact. Individual differences in parameters reflecting rate of processing, or loss of information from queues, may have a cumulative effect on performance. Measures of individual differences at early stages might then be relatively uncorrelated, but measures at later stages might increase in complexity and in intercorrelation. Individuals might differ also in the sequence each pursues through a common set of steps, some taking step three before step two, or further, in the kinds of alternative steps or routes through which information is processed. These possibilities would also make interrelations among parameters complex.

Categories of process differences. To keep these logical possibilities distinct, then, we can define four different forms or sources of individual differences in information processing: parameter differences(p-variables); sequence differences (q-variables); route differences (r-variables); and summation or strategic differences (s-variables). The distinctions between p-, q-, and r-variables can be clarified by imagining two flow charts that characterize the performance of two different individuals on some task; p-variables would refer to differences between the individuals on particular steps or components (e.g., capacity of STM, time needed for stimulus encoding, etc.); q-variables would be shown by the two flow charts taking the same steps but in different sequences (e.g., early vs. late work on some subgoal); r-variables would be indicated by the inclusion of qualitatively different steps in the two flowcharts (e.g., visual

image rotation, or double checking, used in one chart and not in the other.)

This extension, together with the fact that we are ultimately concerned with individual differences in complex learning and problem solving, suggests the need for the fourth category, s-variables, representing individual differences in summative, strategic, or other more molar properties of information processing models. The category of s-variables would include gross differences in the assembly and structure of the program systems used by different subjects (as opposed to r-variables representing variations within the same basic program). Taylor (1976) classified memory models as being of four basic types: serial vs. parallel processing, and exhaustive vs. self-terminating search. Individual differences associated with these kinds of differences in models would be classed as s-variables.

But other variables are implied here as well. Laboratory tasks and test items are often repetitive and noncumulative. They require little or no prior knowledge on the part of the learner. One might imagine a learner cycling through a sequence of steps for each item in a test or list to be learned, without showing intercycle interaction. In contrast, learning from instruction is usually cumulative; prior knowledge, skills, and predispositions are intimately involved in present learning, and instruction is usually geared to take advantage of these. Even repetitive laboratory tasks have their cumulative properties, as when an early guess builds confidence, continued practice yields a stable strategy, or persistent errors breed anxiety, defensiveness, or change in strategy. Learning-to-learn, transfer, and retention differences operate across trials and tasks, and may do so also across items and tests. There is evidence that familiarization alters individual differences in learning (Cronbach & Snow, in press), that learning occurs within tests (Whitcomb & Travers, 1957), that early and late parts of some tests may relate differently to learning task measures (Koran, Snow, & McDonald, 1971), that anxiety is both a predictor and a product of learning task performance (Gaudry & Spielberger, 1971), and that various motivational constructs may moderate the role of ability in learning (Snow, 1976). While simple cognitive models may provide parsimonious starting points, we expect them to be far too simple in their account of individual differences in instructional learning. The category of s-variables keeps this likelihood in focus.

Computer simulations of complex problem solving appear to bear out the importance of individual differences in q-, r-, and s-variables, in addition to p-variables. They further suggest that observed individual

differences are not subsumable under a few general constructs. In their work, Newell and Simon assume a unidimensional homogeneous ability scale for each task environment they analyze, with no assumption that this scale is generalizable to (i.e., correlates with) the scale for another task. They further restrict individual difference variation by making separate task analyses for subjects who appear to use radically different problem spaces for a task, and by selecting them to be homogeneous on key task dimensions; all subjects are accomplished with respect to some task dimensions and naive with respect to others. Yet marked individual differences, as well as commonalities are observed in the problem-solving protocols Newell and Simon collect. Abstracting from the fine details of these protocols, Newell and Simon reach statements about the kinds of individual differences observable in complex problem solving (of the sorts they studied, namely, cryptarithmetic, logic, and chess). We can paraphrase their observations as follows:

1. Subjects differ in the detailed contents of LTM when beginning a problem. This places constraints on the problem spaces and programs available for use.
2. Subjects differ in the way they characterize the initial problem. They learn gradually which aspects of the problem should be given first priority and which can be ignored.
3. Subjects differ in persistence in pursuing a subgoal and, conversely, in their readiness to return to the overall problem in pursuing complete solution.
4. Subjects differ in the priority given to restructuring the problem as information is acquired, as opposed to working in the framework of a definite plan.
5. Subjects differ in the cues used to detect lack of progress toward a goal.
6. Subjects differ in trying to explore paths mentally, as opposed to writing out expressions for examination.
7. Subjects differ in the degree to which operational rules are associated directly with problem features, reducing the need for searches to find appropriate rules.
8. Subjects differ in acquiring knowledge that certain features are not remediable, indicating termination of search, while other goals are automatically achieved by fixed sequences.

These kinds of differences suggest the constructive, strategic, adaptive character of individual differences on complex tasks. Research on individual differences in information processing has been mostly limited to date to a consideration of p-variables on a few simple tasks. But, in the words of Newell and Simon (1972),

> Substantial subject differences exist among programs, which are not simple parametric variations but involve differences of program structure, method, and content. Substantial task differences exist among programs, which also are not simple parametric variations but involve differences of structure and content. (p. 788)

Simon (1976) has gone on to describe several additional examples of ability variations manifested in various computer simulations. He concludes his latest summary with the following statement:

> This description of the processes required for intelligent performance in a half dozen disparate task environments sounds a bit like an argument for Spearman's _g_. A number of basic structures and processes show up again and again in the various programs. If we equate abilities with these kinds of structures and processes, then it appears impossible to construct any simple isomorphism between particular abilities and particular task environments (a result that is consistent, I think, with the experience with factor analysis).
>
> If this be true, why does _g_ account for only a modest part of the total variance in intelligence? First, we have seen that, while there is a great commonality of process, most tasks require also some very specific knowledge (words, perceptual tests, familiar chunks), and expert performance of some tasks calls for an enormous amount of such knowledge. Second, while the same basic processes show up in many different tasks, a given process may be employed more or less frequently in different task environments. Third, the basic processes may be combined in more than one way to produce a program for performing a particular task. In [one] problem, we described four different programs that draw to some extent upon different underlying abilities, and that may differ greatly in effectiveness. Proficiency in a task may depend on how the basic processes and relevant knowledge have been organized into the program for task performance.
>
> Finally, it is not certain to what extent _g_ is to be attributed to common processes among performance programs, or to what extent it derives from individual differences in the efficacy of the learning programs that assemble the performance programs. (p. 96)

Thus the problem for further research will be to distinguish p, q, r, and s sources of individual differences, and to show how they can be combined and/or further differentiated. Information processing models of particular tasks or tests will need to show how these kinds of differences

work in consort to produce observable differences in performance. Aptitude variables (A), and instructional treatment variables (T), and their interactions will need to be analyzed and understood in these common terms. This will best be accomplished by a combination of correlational research that relates A, p, q, r, and s variables, and experimental research that manipulates T in ways that influence these relationships. This suggests again the value of an elaborated S-R-R paradigm.

Psychometrics

Before formulating in more detail a research strategy that seems to emanate from the above discussion, some other methodological points need to be identified and adapted. The earlier discussion of factor analytic models focussed only on substantive and general methodological issues. But psychometric theory has several more specific methods and concepts to contribute to research on aptitude processes as well. These come from two basic tenets of correlational psychology. One is the concept of construct validity (Cronbach, 1971) and its representation in the multitrait-multimethod matrix (Campbell & Fiske, 1959). The other is the concept of dependability of measurement, the machinery for its analysis provided by generalizability theory (Cronbach, *et al.*, 1972), and the use of generalizability estimates in disattenuating theoretical variables.

Generalizability. All psychological measurements are fallible. All psychological theories must stand or fall on psychological measurements, and will be sharpened or dulled by full understanding of the character of these measures. Classical test theory, and the generalizability theory that has superceded it, offer means of analyzing and thereby controlling fallibility in measurement. It therefore gives to theoretical work much more than has usually been appreciated by cognitive psychologists. A generalizability analysis shows whether a dependent variable is capable of detecting significant differences. But it also yields an estimate of theoretical relationship to other measures, as well as a detailed prescription for the conduct of further measurement. Such analyses of test and information processing measures, and of relations between and among them, thus can help explicate features of the constructs they presumably represent.

This is accomplished, first, by partitioning the variance of a measure among its several sources in the facet design of a test, an experiment, or

an experimental task. The analysis of variance components shows the relative importance of different design facets, the generalizability of scores aggregated across various facets, and the number of observations that will be needed for reliable measurement within or between any facets. Estimates of reliability or generalizability can be obtained for any part-score or composite score of theoretical interest. In an experiment, a process measure might well vary in reliability across different experimental conditions. It will be necessary to detect and understand this if sensible theory is to be built upon the experimental results. Related analyses of contrasts in the design can also be used to examine various hypotheses about the independence of process measures presumed to represent independent parts of an underlying model (Calfee, 1976).

Disattenuation. An important additional point here is that generalizability coefficients arising from this kind of analysis allow the relations among process measures to be corrected for attenuation.

Disattenuation has had somewhat of a bad reputation in psychometric work because it tends to paint too rosy a measurement picture in practical situations. Simple reliability coefficients tend to be underestimates, so that corrected validity coefficients are in turn overestimated. Also it is thought, the practitioner by disattenuating lays claim to validity he is not entitled to. In theoretical work, however, primary interest should attach to the "true" relations among theoretical variables, not to relations obscured by fallibility of measurement. For theory, then, disattenuation is indispensible. Relations between ability and process measures or among process measures should routinely be corrected for attenuation. It follows that theory-oriented research on individual differences in aptitude processes will always need good estimates of generalizability of measures.

Construct validity. As noted above, the models currently studied in experimental cognitive psychology are extremely task specific (See also Newell, 1972). A variety of parameters and processing constructs have sprouted from this work, and all could be connected to one or more parts of a general theoretical model. As the tasks embodied in mental tests are also modelled in these ways, still more process measures will be available. While cognitive processes will to some extent always be task-specific, we should be interested primarily in those processes that are

in whole or in part generalizable beyond any one particular laboratory task. At most, this means that a given process must be demonstrable through measurement in (by) alternative independent tasks (methods). At least, a process measure must correlate significantly and sensibly with some other presumed measures of the same or similar process arising from outside of that immediate processing task. One examines this by including in any one study at least two constructs, each measured by at least two methods or tasks. The Campbell-Fiske (1959) reasoning regarding discriminant and convergent validity is then followed. As further demonstration, relations between measures should be manipulable by systematic task or test variations. One proves in these ways that individual differences variance in the measure is to some degree understood. Construct validation is the process of validating a theory (interpretation) of a measure by triangulation.

A Research Strategy

How should the research effort now proceed? Can a detailed strategy be adopted that will not only guide further work, but will also show clearly how and where prior studies connect with an overall framework?

Summary of Starting Points

Theoretical framework. Figure 5 shows the categories of variables that have been identified and indicates by arrows the direction analysis has taken or can take in the future. Standing predictive relationships between aptitude variables (A) and learning outcomes (O) from instruction have been shown to be moderated by instructional treatment variables (T), with the recognition that ATI often occurs. It is clear that AT combinations can be studied in real instructional settings, and should continue to be, but that this research must be supplemented by analyses conducted in laboratory settings where there is more chance of building theoretical models of psychological processes operating in ATI. The cognitive information processing approach of modern experimental psychology seems best equipped to guide and inform such analyses. But computer simulations and related work already completed show that individual differences in these aptitude processes probably take a variety of complex forms. A distinction among four major forms or sources of apparent individual differences in processing should help to unravel these aptitude complexes. It appears that

individuals can differ in parameters (p) reflecting efficiency and capacity in particular processing steps or components, in how a sequence (q) of processing components is organized, in the inclusion of different components or processing routes (r), and in the overall summation (s) of assembly and adaptation of processing to particular tasks. These forms are indicated as p, q, r, and s variables in Figure 5.

- - - - - - - - - - - - - -

Insert Figure 5 about here

- - - - - - - - - - - - - -

Methodology. Research designed to fill in this framework will need to combine multivariate experimental and correlational methods in a general S-R-R paradigm, wherein stimulus conditions are manipulated to test R-R relations by controlling or modifying them. This is not to say that exploratory correlational work will not be useful or that experiments must always include multiple individual measures. But adequate theory will be built and tested primarily on the combined paradigm.

In particular, the requirements of parsimony and construct validation demand the inclusion of multimethod measurement and a representative set of reference factors. And generalizability estimates for all measures will be required, so that key relationships can be disattenuated. More than typical care is also needed in sample selection and description. College populations will differ from high school populations, for example, in a variety of reference factors, whether measured or unmeasured. Attempts to collate findings across studies will need to pay close attention to such "hidden" individual differences.

Procedure

Aptitude selection. One could conceivably start with any aptitude construct of interest practically or theoretically. The main issue of course is whether there is good evidence that the aptitude chosen is related in important ways to learning under given instructional conditions.

An earlier report in this series (Snow, 1976) reviewed the ATI literature on two major aptitude complexes, and these are starting choices particularly to be recommended.

The clearest first choice is G, general mental ability, and the first level of its differentiation into G_c, G_f, and G_v (as shown earlier in Figure 1).

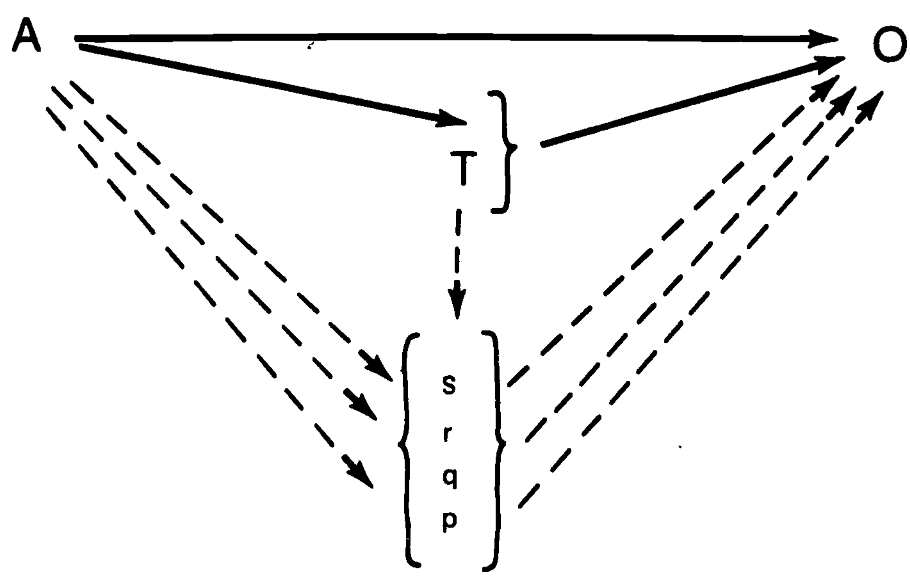


FIG. 5. SCHEMATIC REPRESENTATION OF STANDING CORRELATIONS BETWEEN APTITUDE (A) AND OUTCOME (O) VARIABLES, A x TREATMENT (T) INTERACTIONS, AND THE ANALYSIS OF A AND T VARIABLES INTO INFORMATION PROCESSING VARIABLES. BRACKETS INDICATE COMPLEX INTERACTION.

An information processing analysis would strive to show whether the crystallized-fluid-visualization ability distinction is describable in process terms. A second choice is a motivational complex: achievement via conformity (A_c), achievement via independence (A_i), and anxiety (A_x). These aptitudes have been shown to interact with instructional treatments in their own right, and they appear also to combine with one another and/or with G in higher-order ATI. Information processing analyses that could characterize how such individual differences functioned in combination with G to influence learning would begin to bridge the age-old but artificial gap between cognitive and affective domains, producing a more coordinated and comprehensive view of aptitude.

Reference factors. Any process analysis of an aptitude construct needs to include a battery of representative reference factors. The aptitude constructs identified above would be prime candidates for inclusion in any such battery. In addition, some ability constructs available from differential psychology appear more closely identified with processing concepts and so would justify inclusion. These are: short term memory span, visual memory, perceptual speed, closure speed, and various cognitive style constructs. Measures of these and other ability factors would be chosen from the hierarchical factor model to fit the task at hand.

Exploratory empirical analysis. Some empirical studies have begun identifying individual differences in processes related to learning. (Snow, 1976, began a review of these.) Some aptitude measures have already begun to show correlations with measures from information processing tasks. (One category of these studies is examined in a following section of this report.) We can expect this empirical exploration to continue to generate concepts and relationships bearing on the network depicted in Figure 5. These deserve thorough review in the analysis of any particular aptitude construct.

Task analysis. Given a chosen aptitude, one still needs a fruitful method of task analysis. Different approaches to task analysis have been developed for different purposes, and they represent aptitude-learning relations in different ways. Gagné (1970) constructs learning hierarchies to specify sequences of steps in instructional tasks. One can ask learners to introspect during or after task performance. There are questionnaires that also yield processing protocols. Various experimental arrangements can be made to yield task analytic information. Eye movement records can

sometimes be helpful. Computer simulation methods provide detailed programs and flow diagrams to represent sequences of processing operations in problem solving. And Carroll (1976) as noted earlier has been developing a method that maps ability tests onto a general process model.

At present there is no one best way to gain initial hypotheses about processes in task performance. This is, in many ways, the crux of the whole problem. By one means or another, one seeks a listing of component processes hypothesized to account, in some combination, for individual differences in the task of interest.

Componential analysis. Several aspects of the discussion in previous sections of this paper parallel ideas offered by R. Sternberg (1975). Though developed independently, the two views converge on several common distinctions and emphases. Sternberg has gone on, however, to construct a methodology based on these views which he calls "componential analysis". This seems the best place to summarize his approach. The strategy incorporates many of the principal strengths of information processing, factor analytic, and psychometric models. He has applied it to the analysis of several forms of analogical reasoning problems, of the kind typically found in general mental ability tests. For a full discussion of the approach and application, the reader should consult Sternberg (1975). Only an outline can be presented here.

A componential analysis includes the following steps:

1. Any task or test is composed of items (whether test items or the particular problems used in successive trials in an experiment). Each item is regarded as a composite of subtasks, yielding a composite score for each of a number of individuals. The first step is to identify these subtasks. In Sternberg's work with analogy items, the subtasks are defined by the parts of an analogy. A true-false analogy of the form "A is to B as C is to D" has three subtasks in addition to the total problem process: given an understanding of A, process B-C-D; given A-B, process C-D; given A-B-C, process D.

2. Given an identification of the subtasks, an experiment is designed to obtain what are called "interval" scores for each subtask. In the analogies example, total solution time is regarded as a composite of the solution time needed after A is understood, after A-B is understood, and after A-B-C is understood by the subject. The experiment can be arranged to

provide information in the form of cues separately for each successive part of the task to create these conditions. Four time interval scores (counting total time) are thus obtained. Error as well as latency scores could presumably be used for this purpose.

3. The experimental procedure for breaking up the task item to derive the internal scores is assumed not to have altered the task psychologically. The interval scores are also assumed to be additive, i.e., the composite score contains the interval score for the first subtask, as defined above, which contains the score for the second subtask, etc. This latter assumption is tested by fitting the correlations among scores to the Guttman simplex model.

4. From the interval scores, component scores are estimated using one or more information processing models of the task for each subject. The alternative models are examined in terms of amount of total variance accounted for. The component scores for the best fitting model(s) are adopted as the basic measures of individual differences in information processing for that task.

5. The component model is then internally validated by examining correlations among component scores using the concepts of construct validity and by experiments designed to test implications of the model.

6. The component model is externally validated by examining correlations of component scores, and interval scores, with reference ability factor scores, derived from established reference tests. A structural regression analysis is used to examine the network of relations among component, interval, and factor scores.

7. The previous steps lead to what Sternberg calls an "intensive task analysis" of one task. Such analyses of a series of related tasks then comprises an "extensive task analysis". This aims to demonstrate that the components and models established for one task generalize to account for performance in other, related, tasks.

Thus, Sternberg has taken a major step toward formalizing a procedure that incorporates most of the concepts and methods reviewed in the present paper. Some reservations remain regarding the generality of the procedure, however. While Sternberg applies it to a variety of analogy items, with other tasks the identification of convenient subtasks to obtain a starting

point may not be so readily apparent to the investigator or to the subject. In items from tasks such as the Embedded Figures Test, the Street Gestalt Completion Test, and various spatial ability tests, similar segmentation of subtasks may require considerable experimental work initially. Further, the experimental designs employed to segment analogies and obtain interval scores may not be applicable to all kinds of items, so other experimental designs may have to be invented for some kinds of tasks or tests. Finally, extensive task analysis may be the most difficult step, and this step is not yet well specified. It has been shown by the work of many cognitive psychologists, using whatever method, that process models can be built for single tasks. Combining such models will not simply be a matter of correlating their component scores. General models must be assembled that incorporate all specific models, and these cannot simply be patchworks. Nor can it be assumed that LTM simply contains many specific programs, since this ignores the correlational network of tests and tasks, and in the extreme returns to oversimplified S-R views. And, having strong specific models in hand may in a strange way actually impede progress toward this step, since their accommodation to more general purposes may be awkward, requiring the relaxing of assumptions and much disassembly. Thus, it would appear that the model-builder ought to have an eye on general purposes from the start. This is probably best accomplished by simultaneous analysis of a range of tasks.

But these are mere conjectures before the fact. Sternberg's approach clearly merits extensive trial.

Learning samples. Process analyses of aptitudes need at some point to be brought together with comparable analyses of instructional treatments. A checkstation is required, close to laboratory conditions, where combined assumptions and hypotheses can be examined and perhaps refined. The miniature instructional experiments that have dominated the instructional psychological literature in recent years may be a useful form for this purpose. They are to be regarded as "learning sample tests", however, on a par with the work sample tests of industrial psychology; they probably do not supply direct generalizations to instruction.

Learning samples can also be extracted from real-school instruction, and this may be preferable in many cases. The investigator strives to trace components of his aptitude process model through to individual

differences in actual learning under alternative treatments. Considerable use may be made here also of various forms of task analysis, including learner introspection.

Aptitude test revision. It is likely that conventional aptitude tests will at times need revision to sharpen processing contrasts identified in earlier analyses. Revision may be minor in some cases, extensive in others. The aim is to ensure that the aptitude measure to be used in field research indeed embodies and displays the processing distinctions important in theory. The test is then made a vehicle to ply between field and laboratory.

Instructional studies. The ultimate goal is demonstration that aptitude measures connect to instructional treatment variations in understandable and predictable ways. Aptitudes and instructional treatments described with common process models should make this possible. But the proof is in the real-school, long-duration, instructional studies. Hopefully, laboratory analyses of aptitude and continuing ATI studies in the field will be conducted as parallel, closely-coordinated transactions.

Studies of Short Term Memory: A Demonstration and Critique

No area of experimental cognitive psychology has been treated to more research attention in recent years than that of immediate memory. This is reasonable, since experiments on immediate memory gave birth to the history of research on learning, and since it is likely that immediate memory operations provide a foundation on which deeper processing activities are built. Differential psychologists, too, have long made a place for a "memory span" factor in taxonomies of human ability and have often used paired-associate tests as well. The tests used to represent such factors have been the only cognitive tasks in common use in both psychologies until recently.

Short term memory is thus a natural site for initial exploration of relations between cognitive process measures and cognitive test scores. It is a good place to demonstrate application of some of the points explicated in previous sections of this report. But a word of caution is in order: although immediate memory may be basic to cognitive functioning, it is unlikely to be comprehensive in its yield of concepts of

value in theorizing about individual differences in cognition. Aptitude constructs are unlikely to reduce to memory functions, so a broad perspective needs to be maintained.

Ever since Galton and J. McK. Cattell, there have been sporadic investigations of individual differences in short-term learning and memory tasks and the degree to which such differences were correlated with mental tests. Among the more programmatic efforts were those of Woodrow (1949) and of a group headed by Gulliksen (1942, 1960; see also Stake 1961; Allison, 1960; Duncanson, 1964). Although these authors reported little relation among tests and learning tasks, their results actually support the view that general ability and simple learning correlate. (For discussion, see Cronbach & Snow, in press.) The early findings fail, however, to give a more detailed picture of the bases for such correlations.

The Hunt Studies

E. L. Hunt's program of research relating ability variables to speed and sequence of immediate memory processing has reopened this issue and raised many new questions. The first study reported in the Hunt program (Hunt, Frost, & Lunneborg, 1973) examined the performance of 40 college students on a continuous paired associate task. This task provided estimates of the four parameters identified earlier as defined in the work of Atkinson, et al. (1967). Since Hunt's subjects represented extreme groups on Verbal (V) and Quantitative (Q) ability (as measured by college admissions tests), it was possible to compare means on each parameter among groups of students labelled "high" or "low" on each ability. Herein lie two initial problems that will confuse further research unless recognized.

By convention, extreme groups blocked by partitioning an aptitude continuum are described by terms such as "high", "medium", and "low". But these are absolute terms applied to groups defined on a relative standard. Hunt's "Lows" may turn out to be equal to another researcher's "Highs", or vice versa. At the least, studies using extreme groups designs will need to report normative ranges on the tests used, to qualify the abbreviated labels. In the case of Hunt's subjects, whose admission to college was in part based on these tests, we should prefer the term "medium" as probably a more accurate description for his "low" groups.

A more serious difficulty concerns the meaning of V and Q. Extreme groups formed on the basis of V and Q distributions differ also on all personal variables correlated with V and Q (but left unmeasured in the study). As mentioned previously, there is likely to be a whole network of correlates for any individual difference distribution. Interpretation of results in terms of V and Q is thus tenuous at best. A difference in some information processing parameter associated with V might be associated also with differences in G, verbal fluency, or achievement motivation, or sex, or even in Q itself. In less selected populations, for instance, V will tend to correlate with Q and show mean differences favoring females. This difficulty becomes evident in another way below.

The Hunt analysis concluded that differences on two parameters α and τ were associated with Q. No statistically significant differences were found for the other two parameters (r and θ) or for any parameters on the V contrast. Apparently, students with high quantitative ability showed a higher probability of placing items in STM (α) and a lower rate of loss of information from ITM (τ) than did students scoring at a medium level of quantitative ability.

However, a plot of the reported means, in Figure 6, suggests a different story, and this underscores several of the methodological preferences stated previously. There are three implications to be drawn from the figure, none of which were apparent in the statistical tests applied to means or given attention in the authors' interpretation.

- - - - - - - - - - - - - -

Insert Figure 6 about here

- - - - - - - - - - - - - -

1. Both V and Q ability appear to relate to both parameters, though admittedly the relation of Q is clearly the stronger. This implies that a more general ability, not Q alone, underlies the correlation with both parameters. Considering that V and Q are probably correlated in the population, and taking the hierarchical view of ability organization proposed in Figure 1, both abilities are subordinate to G_c and G. The principle of parsimony demands that special ability interpretations be adopted only when general abilities can be ruled out. They cannot be here, because no general ability measures were included. One could have been approximated, however, by forming two composites, V+Q and V-Q. The first

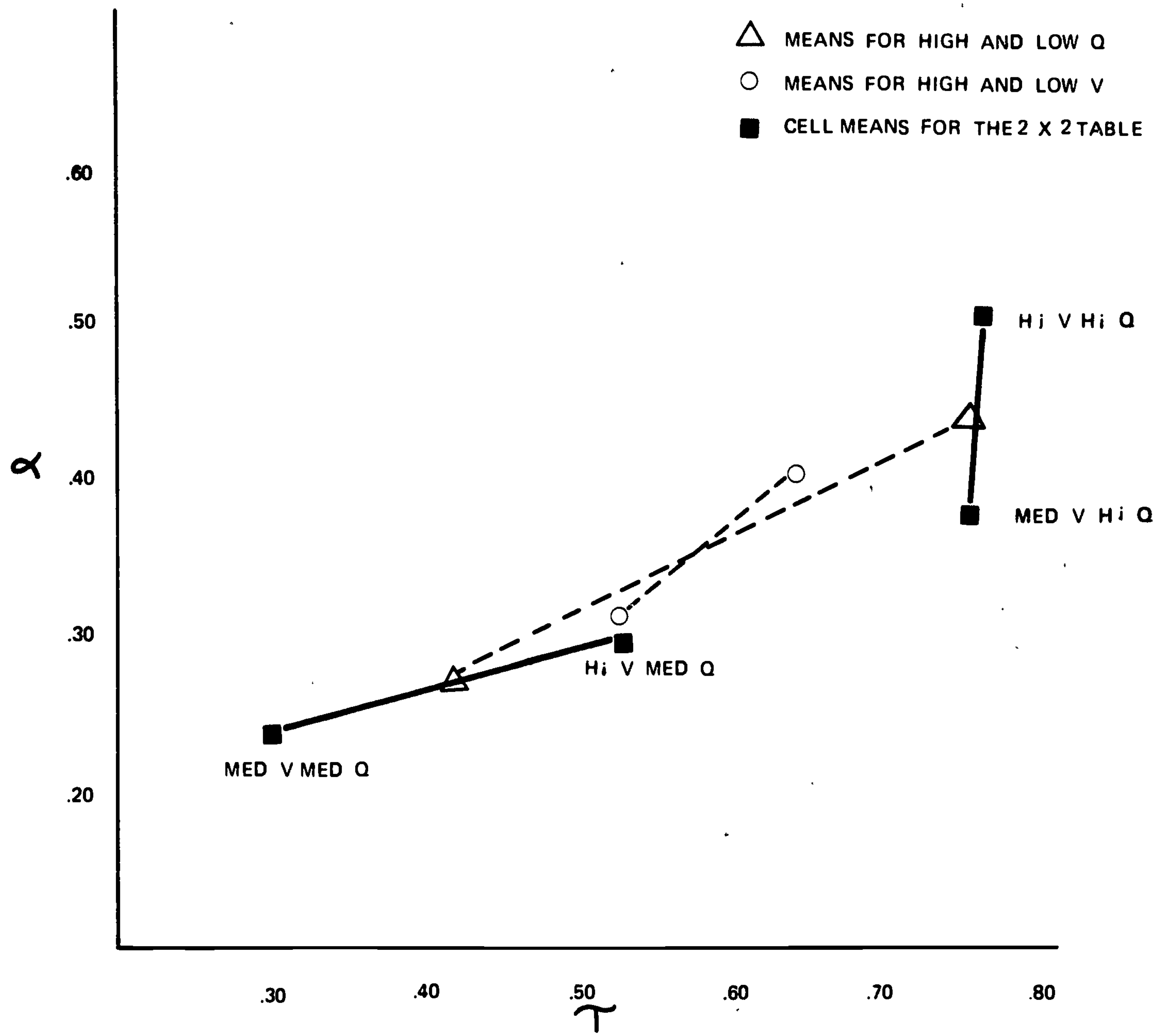


FIG. 6. PLOT OF MEANS ON TWO PARAMETERS (α AND τ) OF THE ATKINSON - SHIFFRIN MODEL FOR GROUPS OF SUBJECTS DIFFERING IN VERBAL (V) AND QUANTITATIVE (Q) ABILITY (AFTER HUNT, FROST, AND LUNNEBORG, 1973).

composite then represents G, while the second gives an independent linear contrast to test the special ability hypothesis. Normally this is best done in continuous distributions rather than in extreme groups, but it could have clarified interpretation here.

2. If some general ability relates to both parameters, we should expect the two parameters to be intercorrelated. If one approximates the G continuum by attending only to the High-High and Medium-Medium groups (the general dimension), the plot clearly suggests this, and Hunt, et al. reported a correlation of 0.42 between α and τ later in their chapter. Unfortunately, the matrix of intercorrelations did not include the ability tests, so the relative strength of these various relations could not be studied.

3. In the plot, Q appears to moderate the relation between V and the two parameters. Among high quantitative students, V relates to α. The vertical line is nearly parallel to the ordinate, suggesting that this correlation is almost perfect, at least when means rather than individual scores are examined. Among medium quantitative students, V relates to τ. The mean slope here approaches the horizontal, again suggesting high positive correlation. In the Hunt analysis, row and column means were tested for statistical significance; apparently this moderator relationship (an interaction) was not. These relations may be weak, so conclusions are unwarranted. But overly conservative statistical analyses, which minimize Type I errors at the expense of Type II errors, are also unwarranted at this stage of research. This moderator trend is an example of the subtleties with which one must deal in combining experimental and correlational methods and concepts. Such a finding deserves to be kept as an hypothesis for further research, not swept aside by the insensitivities of conventional modes of hypothesis testing. If confirmed, the hypothesis would state that verbal ability is associated with entry of information into STM among learners of high quantitative ability, but is associated more with the rate of loss of items from ITM among learners middling in quantitative ability. It is noted here that the task involved largely numerical responses and that floor and ceiling effects, ubiquitous problems in correlational research, cannot be ruled out.

Another small experiment in the Hunt series compared ability groups on reaction time for matching letters by physical identity, or by name when different in physical shape. It was found that high V students were substantially faster than medium V students in name matching, but not in physical stimulus matching. This would imply that verbal ability is associated with speed of coding linguistic stimuli. Here again, however, means for Q showed a similar but nonsignificant trend, suggesting an underlying relationship with a more general ability.

Two other Hunt experiments reported results only for the V contrast. One used a task in which proactive inhibition was built up over a series of trials; vegetable names had to be held in memory while the subject counted backward. Degree of release from proactive inhibition was then measured on a trial in which the words to be recalled switched from vegetables to occupations. Both high and medium V subjects showed the release effect when number of words recalled served as the measure. But high V subjects showed markedly greater release when performance was scored for number of words recalled in the correct serial order. In other words, temporal encoding appears to deteriorate in middle V learners, while high V learners maintain temporal order. This is consistent with other data the authors reported. It was noted that temporal coding and speed of coding are closely related. The other study used Sternberg's (1969) STM search task, where the subject must determine whether a stimulus digit is or is not one of a set held in memory from a previous presentation. Reaction time scores for memory sets of differing size yielded a slope parameter reflecting speed of search. High V subjects showed faster search speeds than did medium V subjects.

Summarizing these and related studies, Hunt et al. drew the conclusion that verbal ability is associated with speed of coding, temporal-order preserving, and search operations in STM. Given the uncertainties introduced by the V and Q extreme groups design, a more likely hypothesis for further work might be that differences in these processes relate to G, at least within its middle to high range.

The Seibert-Snow Studies

It is appropriate now to discuss some earlier research conducted in the traditional correlational mold. This will show some of the powers and also some of the weaknesses of such studies, relative to the Hunt approach. Two studies (Seibert & Snow, 1965; Snow & Seibert, 1966; and Seibert, Reid, & Snow, 1967) were designed to explore individual differences in some temporal features of initial stimulus processing in the visual system and their relation to tested abilities. Their general purpose was to test and extend the Guilford taxonomy by demonstrating that motion picture tests tapped abilities that could not be fitted into existing categories. It was reasoned that the factor analytic conception of human intellect was based almost solely on printed tests. The dynamic character of film, and the control of temporal and spatial features of stimuli that it provided, would allow not only a more comprehensive survey of human abilities but one that might come closer to process descriptions of ability than traditional measurements via printed media. In particular, motion pictures might be used to present experimental stimulus arrangements as group tests.

A large number of motion picture tests were constructed. These were combined with several existing motion picture tests available from wartime work in the Aviation Psychology Program (see Gibson, 1947), as well as a battery of printed reference tests, and investigated in a series of exploratory factor analyses. While some specific hypotheses were built into the tests, these were not precisely stated or tested in the exploratory phase (and as it turned out, the only phase) of the research. Discussion will be limited here to one particular exploratory hypothesis.

Several motion picture tests were constructed to allow group test administration of experimental conditions like those used by Averbach and Coriell (1961) to study the "erasure" or visual backward-masking phenomenon. They had followed Sperling's (1960) early work on information processing hypotheses in the visual system, but had not noticed individual differences among their three subjects. We thought that there might well be important individual differences in the effects of backward masking, and that different abilities might relate to performance at different points in a backward masking curve. The test was designed to present

randomly-constructed letter arrays tachistoscopically. In each item, one letter of the array was marked by an adjacent bar or circle appearing simultaneously or at a short interval later. Eight delay intervals, ranging up to 510 msecs., were included with each item incorporating one delay interval. A score for each delay interval showed the number of marked letters correctly recorded on items at that interval. Following Averbach and Coriell, the condition in which the letter was marked simultaneously by a bar marker and after an interval by a surrounding circle, should give the effect called "erasure".

A sample of 100 male freshmen engineering students at Purdue University served as subjects. With a delay interval of 94 msec., average performance was quite inaccurate. At shorter or longer delay intervals, average performance was relatively accurate. These findings replicate the results of Averbach and Coriell. But there were marked individual differences at each delay interval. Correlating these differences with ability factors drawn from a factor analysis of other motion picture and conventional tests, it was possible to project ability factors into the delay interval space to show how much variance in performance could be accounted for by each ability factor. It was found that subjects with high scores on an ability factor called "perceptual integration" (PI) did well at intervals shorter than 94 msecs.; subjects with low scores on this factor did poorly. The factor represented tests requiring identification of pictures of common objects presented on the screen, with parts masked and/or with parts distributed as bursts, over a short time span. Measures of closure speed and visual acuity also correlated with this factor. Subjects high on a factor called "verbal facility" (V), including conventional vocabulary tests as well as measures of tachistoscopic letter and word recognition, did well at intervals longer than 94 msecs.; those scoring lower on this factor did poorly here. Figure 7 shows the percent of variance accounted for by these factors at each delay interval. It also suggests that a third ability factor, perceptual speed (PS), may be relevant to performance at about 260 msec. delay, and that PI may become relevant again at later delay intervals. Categorizing subjects as high-high, high-low, low-high, or low-low on the PI and V factors, respectively, yielded the four average curves shown in Figure 8.

- - - - - - - - - - - - - - - -

Insert Figures 7 & 8 about here

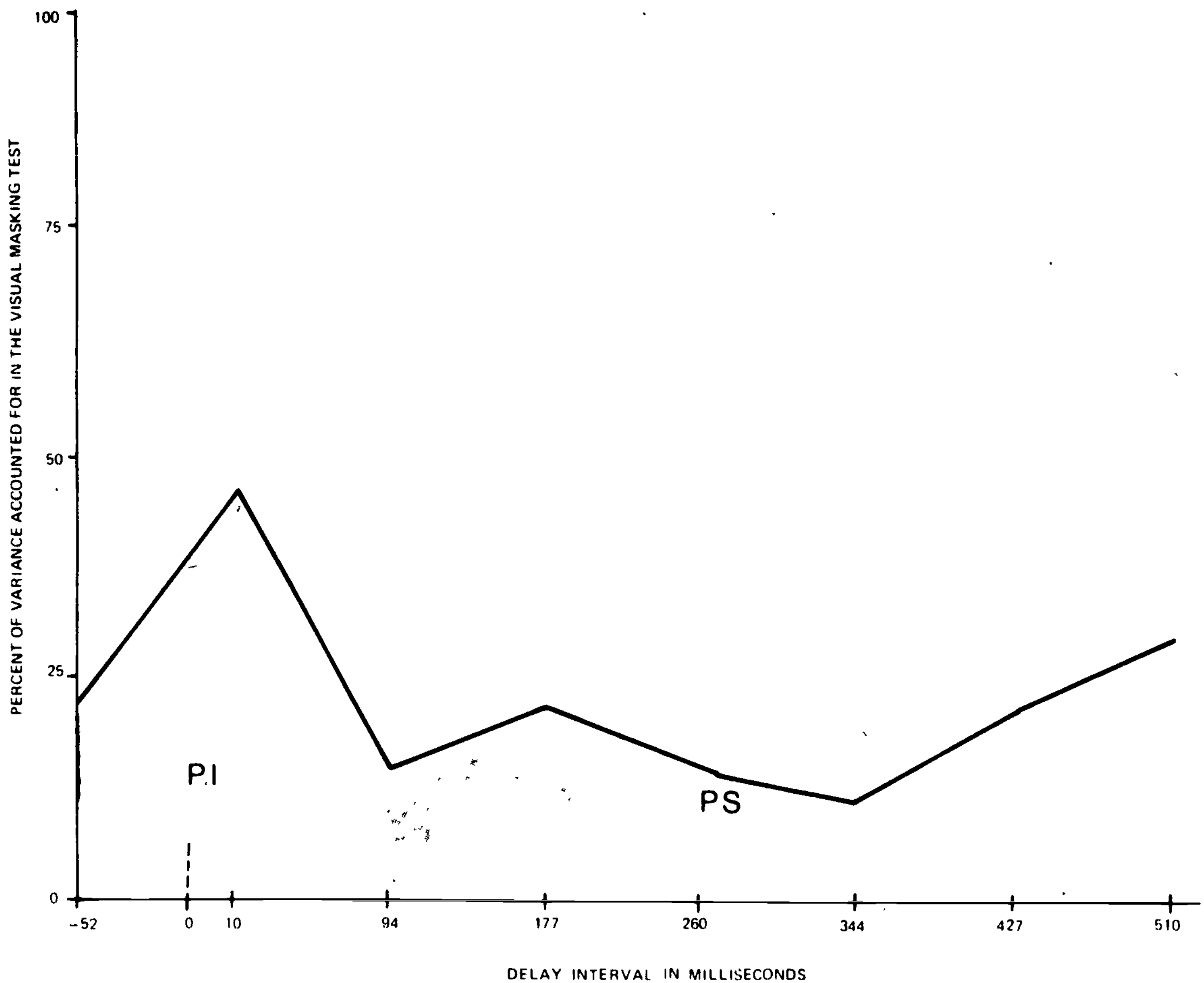


FIG 7 PERCENT OF VARIANCE ACCOUNTED FOR AT EACH DELAY INTERVAL IN THE VISUAL MASKING TEST BY THE PERCEPTUAL INTEGRATION (PI), VERBAL (V), AND PERCEPTUAL SPEED (PS) ABILITY FACTORS N= 100 PURDUE ENGINEERING STUDENTS

48

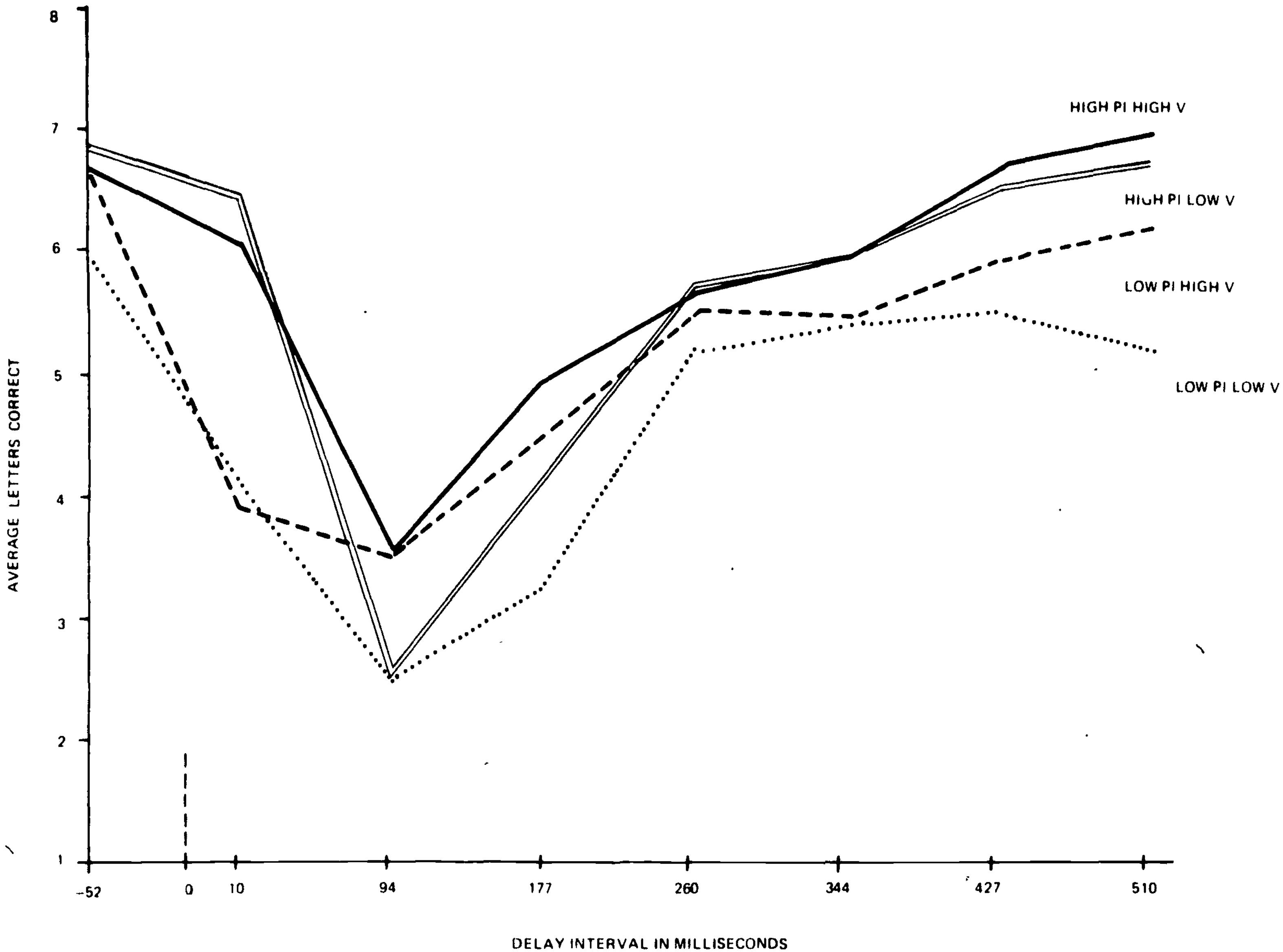


FIG 8 AVERAGE NUMBER OF LETTERS CORRECT AT EACH DELAY INTERVAL FOR GROUPS OF STUDENTS (N=25) DEFINED AS HIGH PI AND HIGH V, HIGH PI AND LOW V, LOW PI AND HIGH V, AND LOW PI AND LOW V

Since the subjects were freshman engineers, they were probably more similar to the high Q than to the medium Q subjects of the Hunt et al. studies. The verbal factor, however, was not defined by a college admissions test, and included some STM measures in addition to vocabulary. It might be reasonable to equate Hunt's high V subjects with the high V subjects of the Seibert-Snow research. But a similar equation for medium vs. low subjects in the two studies might not be safe. The two ends of the Seibert-Snow perceptual factor are fairly described as "high" and "low" since engineering students are probably unselected on such a variable.

Treating abilities as continua in multidimensional space is superior as an exploratory strategy to the extreme groups procedure used by Hunt. With the multivariate analysis completed, one can then always produce mean curves of the sort shown in Figure 8, if these are considered more descriptive or understandable than the "geological", percent of variance, chart of Figure 7. In another respect, however, the Seibert-Snow study suffers from a problem exactly like that of Hunt. The factor analysis was conceived in the Guilford tradition, and relied on varimax rotation to define a long list of special ability factors. Hence, no general ability construct was defined; so one cannot be sure that V is the key ability here as opposed to G or G_c. The perceptual factors, PI and PS, suffer similarly from uncertainty of interpretation.

The possibility that a more general ability may be implicated is apparent in Figure 9. This shows the overall mean curves on the backward masking task for three samples of college students. The 100 Purdue engineering students used in the previous analyses are shown as a solid curve. Another 159 subjects from a later study were relatively unselected Purdue undergraduates in humanities and sciences. Their average curve is shown as a blank-core curve. Recent data on the same task has also been collected from 25 Stanford undergraduates; their average is the small-dash curve. (For details on this study, see Snow, Marshalek, and Lohman, 1976.) It is likely that these three samples differ on general scholastic ability. Certainly they differ in the selectivity exercised at admission by their respective academic programs. These differences correspond to the differences in elevation of the curves shown in Figure 9, and can be seen on both sides of the 94 msec. point. In short, general ability cannot be ruled out as the basic correlate of performance on this task.

- - - - - - - - - - - - - -

Insert Figure 9 about here

- - - - - - - - - - - - - -

Other problems plague virtually all correlational surveys of this sort. Test administration is typically conducted with the group of subjects together in a large auditorium. Order of test administration is the same for all subjects, and interaction among subjects, while controlled, can never be ruled out altogether. Order effects, and sources of error common to tests adjacent in time,can influence the pattern of correlations obtained. In addition, studies that present experimental conditions on film under such conditions must deal with the confounding effects of viewing-distance, viewing-angle, and visual acuity. Seibert and Snow assigned subjects to seats randomly and included an individual administration of a visual acuity test. These variables could then be included in the analysis and controlled statistically. The PI factor discussed above included individual differences in viewing distance and visual acuity in its definition. All these problems lead to doubts about adequacy of controls and replicability of conditions.

There is, on the other hand, a benefit to be derived from the relative lack of controls in correlational as opposed to experimental studies. Figure 9 displays curves obtained from three different large group administrations of a film designed to approximate the exacting experimental conditions applied by Averbach and Coriell (1961) to three laboratory subjects. The form of the curves is similar enough to that obtained by Averbach and Coriell to suggest that the phenomenon of backward masking is generalizable over substantial variations in task and administration conditions, and subject samples.

Despite the limitations of the Seibert-Snow research as an exploratory correlational survey, the findings do lead to hypotheses worth further test, and these connect loosely to the Hunt, et al. results. The implication of Figures 7, 8, and 9 is that subjects differ in the severity with which the backward masking effect occurs and the temporal location at which it occurs. These differences are associated with differences on tests of perceptual and verbal abilities. Perhaps masking at about 100 msec. delay marks the point where in the average subject information is being transferred from a sensory buffering operation designed to register and accumulate bits of percept to one designed to code sums of these in STM.

ERIC

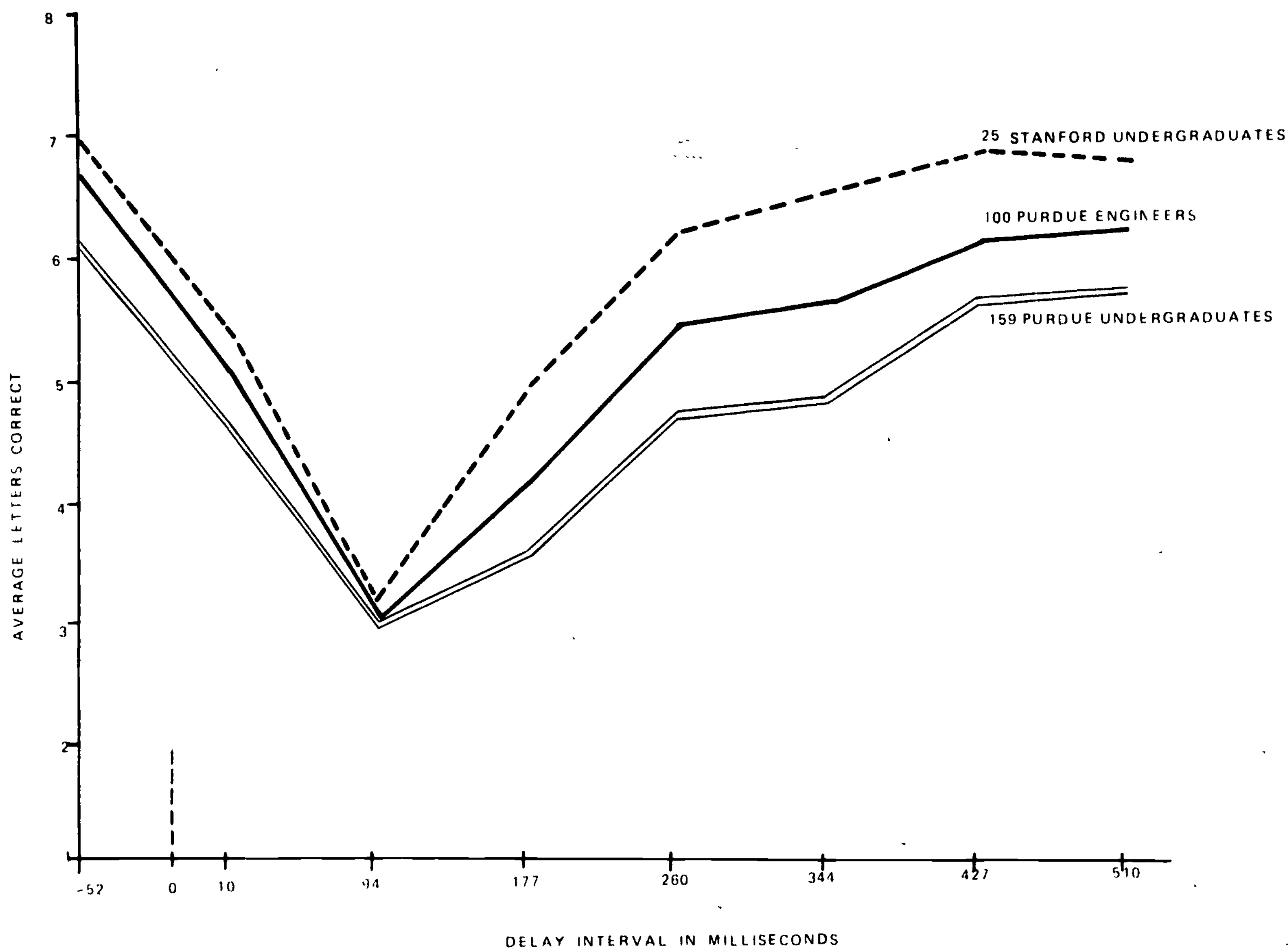


FIG. 9. AVERAGE NUMBER OF LETTERS CORRECT AT EACH DELAY INTERVAL FOR GROUPS OF STANFORD UNDERGRADUATES (N 25), PURDUE ENGINEERING UNDERGRADUATES (N 100), AND PURDUE UNDERGRADUATES (N 159)

This would associate the result with the first Hunt finding relating parameter α, the probability of entry of an item into STM, to V among high Q subjects (if the engineering students used by Seibert & Snow are assumed to be high Q). But there are two alternative interpretations for such a finding. Both can be understood within a three-component model consisting of 1) an iconic sensory buffer or store, 2) a symbolic recognition-encoding mechanism, and 3) STM.

One view would hold that individual differences arise from two p-variables: differences in strength (or resistance to decay) of the initial iconic image, reflected by the PI score; and differences in speed of encoding into STM, reflected by the V score, assuming Hunt's hypothesis is correct. Then, we assume that the masking stimulus causes erasure only if it occurs during transfer (coding) of information from iconic to STM store; it is perceived as superimposed on the array if it occurs while the array is in iconic store, and is coded by array position if it occurs after the array is coded into STM. Thus, subjects who are high V (and high Q in this sample) are able to code letters into STM more rapidly than are low V subjects, so erasure occurs at shorter delay intervals for them. Subjects who are high PI retain a more lasting iconic image on which their encoding processes can work, so erasure occurs later for them. Encoding, and erasure, then would occur early if a subject is high V-low PI and later if a subject is low V-high PI. The high-high subject has the benefit of both lasting iconic images and fast encoding. He would show the best overall performance. The low-low subject's performance would show more pronounced erasure effect through a wider range of delay intervals. Figure 8 seems to bear this out.

An alternative view posits a single q-variable as the source of process differences. Subjects high in V code the array symbolically into STM from iconic store, and the masking effect operates as before. However, subjects high in PI transfer the iconic image directly to STM as an image; symbolic encoding occurs at output from STM during response, rather than at input to STM. The masking effect still operates on the symbolic encoding process, but it comes later for these subjects because the encoding step comes third rather than second in the model. Perhaps the masking stimulus "catches up" to the array image in STM, interfering only if encoding has started.

ERIC

Thus, this alternative hypothesis, though less plausible than the first in today's parlance, is more parsimonious with respect to individual differences; it posits a single process difference concerned with the order in which a series of three processing steps are carried out, and suggests that the aptitude score difference, V-PI, would index that process difference. It avoids hypothesizing individual differences in iconic image decay and speed of encoding, or associating these directly with separate aptitudes. The two alternatives are depicted schematically in Figure 10. Further research will need to contrast these alternatives experimentally, and to distinguish V and PI from G correlationally.

- - - - - - - - - - - - - - -

Insert Figure 10 about here

- - - - - - - - - - - - - - -

The masking or erasure effect appears to be useful in probing for individual differences in the initial visual processing system. Research on visual masking has grown substantially in the past decade, and there are now sophisticated cognitive, as well as sensory physiological explanations for the phenomenon. (See Kahneman, 1968; Turvey, 1973; and Breitmeyer & Ganz, 1976.) But individual differences have not been considered in this work.

It was noted above that one of Hunt's studies showed V (or G) related to speed of search in Sternberg's (1969) STM task; that is, the individual slopes relating memory set size to reaction time (RT) for each subject were shallower among more able subjects and steeper for less able subjects. This seems also consistent with the Seibert-Snow data. Each 8-letter array is a memory set. The subject must somehow fix as much of it in STM as fast as possible, to be able to match the marker to one of the letters when it appears. Averaging across delay intervals, the high ability subjects (high G?) clearly do best. This in turn leads into further research by Chiang and Atkinson (1976).

The Chiang-Atkinson Study

Chiang and Atkinson administered the Sternberg memory search task, a visual search task (Neisser, 1967; Atkinson, Homgren, & Juola, 1969), and a standard digit span task to a sample of college students, of which half were male and half female. They also collected SAT-V and SAT-Q test scores from university records.

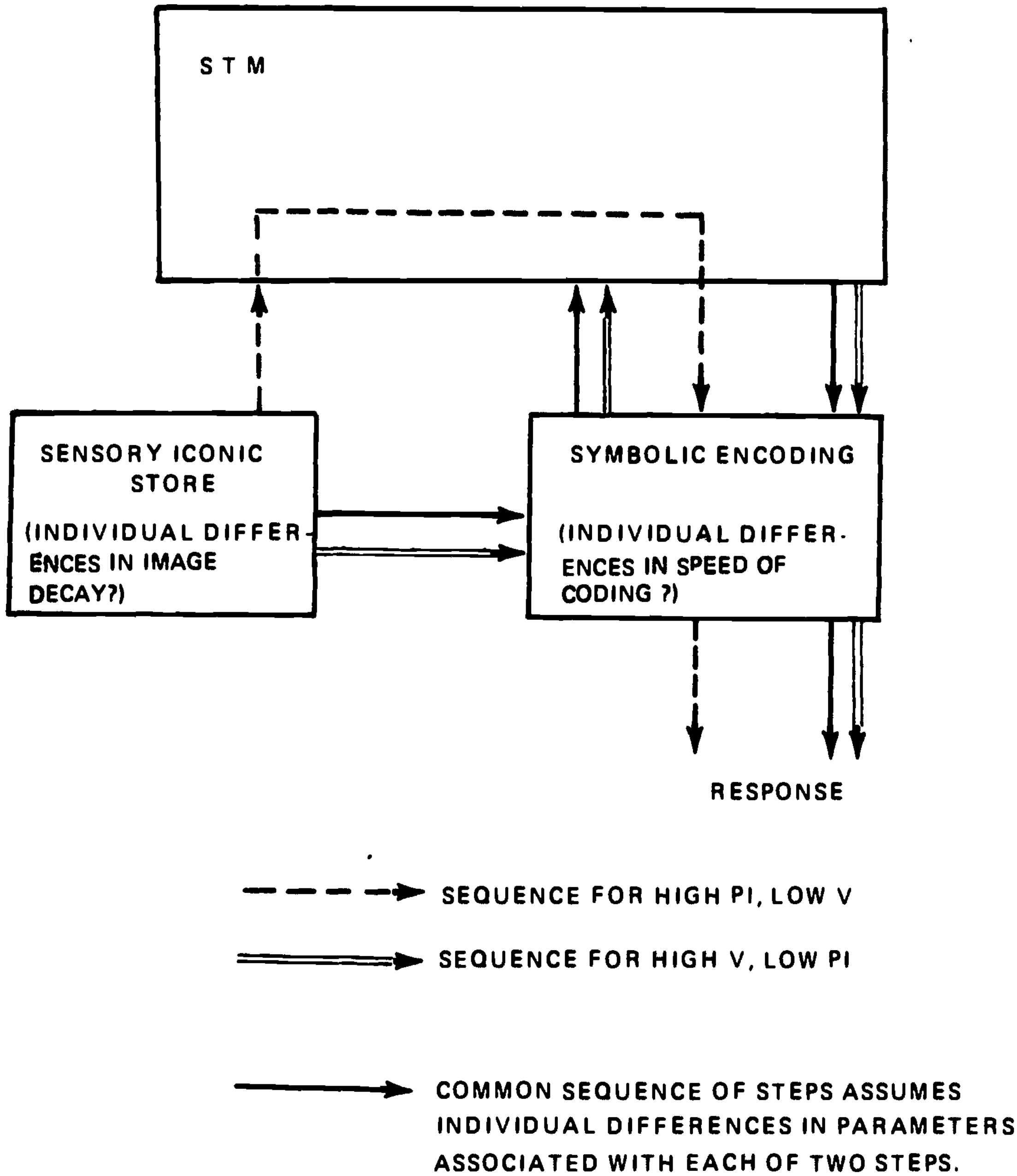


FIG. 10. SCHEMATIC REPRESENTATION OF HYPOTHESIZED INDIVIDUAL DIFFERENCES IN TWO P-VARIABLE vs. ONE Q - VARIABLE TO ACCOUNT FOR VISUAL MASKING RESULTS.

There are two points of interest here. First, because the Chiang-Atkinson design included two search tasks, and assessed reliability for each of the measures derived from these tasks, it is possible to conduct a more penetrating correlational analysis of their data than is typical in experiments on STM. We can apply the machinery of test theory to examine the construct validity of the measures and the theoretical models underlying them. The second concern is whether the Chiang-Atkinson data replicate those of Hunt. Chiang and Atkinson pushed their data analysis further than many experimenters typically do, with surprising results.

Table 2 reproduces the correlational data. The reliabilities shown in the main diagonal are split-half estimates corrected using the Spearman Brown formula. Chiang and Atkinson also computed test-retest reliabilities across experimental sessions. These are not reproduced here. As shown, reliability was quite adequate for all parameters.

- - - - - - - - - - - - -

Insert Table 2 about here

- - - - - - - - - - - - -

High correlations were obtained between the two slope parameters, and also between the two intercept parameters. These plus the relatively low correlations between the within-task pairs of slope and intercept parameters constitute evidence of construct validity following the traditional multitrait-multimethod reasoning of Campbell and Fiske (1959). In brief, this is that measures representing the same trait but based on different methods of measurement should correlate more highly than measures of different traits based on the same method of measurement. The very high correlations among like parameters here allowed Chiang and Atkinson to form two composite scores for overall slope and intercept in their later analyses. However, the like parameters are thought to represent different components in the information processing models of each task. We can pursue the construct validity question further, using the individual correlations to test the basic models.

The reliability estimates allow us to correct the observed parameter intercorrelations for attenuation and thus to see what correlations should obtain in theory among the true parameter scores. The correction formula is simply

$$r_{T_x T_y} = \frac{r_{xy}}{\sqrt{r_{xx}}\,\sqrt{r_{yy}}}$$

TABLE 2

Correlations and Reliability Estimates Reported by Chiang and Atkinson (1975; N=30)

Variable	MINT	MSLOPE	VINT	VSLOPE	MSPAN	SAT-V	SAT-Q
MINT	(.96)	.11	.97*	.43	-.33	.20	-.39
MSLOPE		(.89)	.04	.83*	.04	.19	.03
VINT			(.95)	-.29	-.35	.14	.29
VSLOPE				(.91)	-.08	.34	-.05
MSPAN					(.95)	-.18	-.09
SAT-V						---	.44*
SAT-Q							---

Where: MINT=Memory Search Intercept, MSLOPE=Memory Search Slope, VINT=Visual Search Intercept, VSLOPE=Visual Search Slope, MSPAN=Digit Span Average, SAT-V=Scholastic Aptitude Test Verbal Score, SAT-Q=Scholastic Aptitude Test Quantitative Score. Split Half Reliability, appearing in the main diagonal, was estimated by the correlation of scores on Blocks 1, 3, 5 with scores on Blocks 2, 4, 6, stepped up by application of the Spearman Brown Formula.

* = Significantly different from zero at .05 level

where: $r_{T_xT_y}$ is the true score correlation between two measures

r_{xy} is the obtained correlation

r_{xx} and r_{yy} are the reliabilities of the two measures.

Using the test-retest reliabilities reported by Chiang and Atkinson in this formula, the theoretical correlations between like parameters exceed 1.00, suggesting that the test-retest estimates were too low. This is likely as they were based on intersession correlations that displayed practice effects, and Chiang and Atkinson appear not to have adjusted these coefficients by Spearman-Brown as they did the split-half coefficients. Using these latter estimates (See Table 2), the intercept intercorrelation becomes 1.00 while the slope intercorrelation becomes .90. These results suggest that the relation between slopes and also between intercepts across the two tasks is close to perfect. According to the underlying models, however, the two slopes and the two intercepts should not reflect exactly the same information processing components. Further, the observed correlation in Table 2 between VSLOPE and MINT is 0.43, not to be ignored, though of borderline significance statistically. When corrected for attenuation, this correlation becomes 0.45.

These results are examined further in Figure 11, which shows the presumed model components for each parameter, following Chiang and Atkinson, together with the theoretical correlations obtained by correcting for attenuation. Within-task correlations have not been corrected or shown in the figure, since errors of measurement within the same task are not independent. Note first that the two intercept parameters differ by one component, yet correlate perfectly. Such a result might arise for one of three possible reasons: 1) Stimulus encoding might be a constant, indicating that there are no individual differences in this stage of processing. 2) Individual differences in stimulus encoding might be perfectly coincident with individual differences in binary decision and/or response production and thus deserve no status as a separate ability construct. 3) The models of one or both tasks might be wrong in some way. If one assumes for the moment that the models are correct, then the three other correlations make sense. VINT and MSLOPE have no process components in common and correlate near zero. MINT and VSLOPE have one process component in common and show a moderate correlation, representing about 20% common

variance. But this common component is stimulus encoding! If the intercorrelation arises from individual differences in stimulus encoding, then the first alternative explanation for the results listed above must be incorrect; stimulus encoding cannot be a constant. This produces doubt about the second alternative explanation as well, even though it cannot be logically eliminated. There is in addition other evidence that individual differences in stimulus encoding can be measured reliably and correlated significantly across independent tasks (See Ward, 1973, below). We are left with the conclusion that one or the other model must be changed. The only change consistent with the data is to say that individual differences in stimulus encoding are present also in VINT. This makes the perfect correlation between VINT and MINT reasonable while preserving the reasoning associated with the zero correlation obtained between VINT and MSLOPE. Note finally, that the earlier estimate that 20% of the variance in MINT and VSLOPE was due to stimulus encoding is corroborated by the correlation of 0.90 between the two slope parameters; this correlation indicates that 81% of the variance of these parameters is due to individual differences in the comparison stage, and thus 19% is due to stimulus encoding.

- - - - - - - - - - - - - - -

Insert Figure 11 about here

- - - - - - - - - - - - - - -

A second interest is the hypothesis derived from Hunt that the slope parameter should correlate with V. Table 2 shows nonsignificant correlations between the slope parameters and SAT-V, though the correlation for VSLOPE is high enough to be of marginal interest. Also, the correlation of SAT-V and SAT-Q is significantly positive, suggesting that the hierarchical model is correct; a raw V measure is not an isolate. Thus, the data seem to fail here. But Chiang and Atkinson took their analysis a step further. Combining the two slope parameters and also the two intercept parameters, they computed all the correlations separately for males and females. A striking sex difference emerged. Correlations between SAT-V and the combined slope were 0.72 for females and -0.36 for males. A similar pattern was noted for SAT-Q, though the respective correlations were 0.33 and -0.44, somewhat lower. Figure 12 reproduces the Chiang-Atkinson regression plots corresponding to these correlations. MSPAN also gave

PROCESS MODEL COMPONENT

PARAMETER MEASURES	STIMULUS ENCODING	SINGLE COMPARISON	BINARY DECISION	RESPONSE PRODUCTION
MINT	+		+	+
VINT			+	+
MSLOPE		+		
VSLOPE	+	+		

1.00

.45

.04

.90

FIG. 11. TABULATION OF PROCESS MODEL COMPONENTS PRESUMED TO BE REFLECTED BY EACH PARAMETER, WITH THEORETICAL CORRELATIONS BETWEEN THEM.

ERIC

correlations for males of -0.54 and 0.56 with intercept and slope parameters, respectively. The respective correlations for females were -0.02 and -0.24.

- - - - - - - - - - - - - - - -

Insert Figure 12 about here

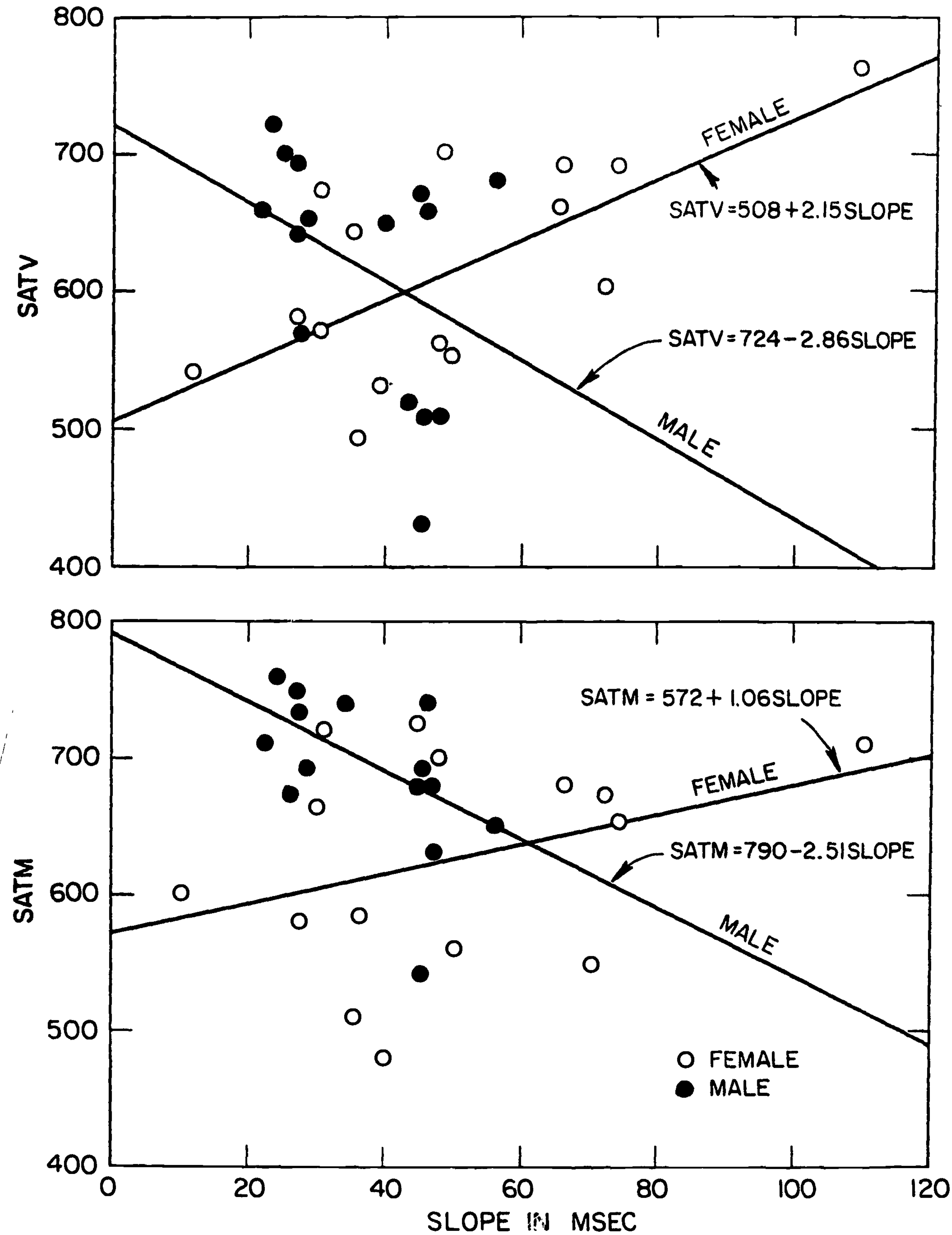
- - - - - - - - - - - - - - - -

Hence, the results are consistent with Hunt's for males only: the higher the male verbal ability score the lower the slope parameter (i.e., the less difference in RT between smaller and larger memory or display sets). But this same statement applies with respect to quantitative ability; SAT-V and SAT-Q are correlated about as highly with one another as either is with the parameters, and none of the correlations are extremely large. If one extracted a general factor combining V and Q, it would probably account for the result.

The opposite pattern appears for females: the higher the verbal ability score (and to a somewhat lesser extent quantitative ability also), the higher the slope parameter (i.e., the more difference in RT between smaller and larger memory or display sets). The findings for females are somewhat uncertain, however, since inspection of the scatterplots shows one and perhaps two female outliers who may have had a marked effect on the regression lines. Removing one or both of these subjects would lower the female correlations but would not change the overall trend of results. (See Snow 1976 for other notes on the significance of outliers.)

Regarding the results for MSPAN, it appears that higher span scores are associated with lower intercept scores and steeper slopes in males but with higher intercept scores and shallower slopes in females. Using the underlying model for interpretation, males appear to be faster at stimulus encoding, binary decision, and response production if they have high memory span scores and faster at single stimulus comparison if they have high V and/or Q scores. Females appear to be slower at single stimulus comparison if they have high V and/or Q scores. Memory span for females shows a slight relation to speed at single comparisons in this same direction, but is unrelated to speed in stimulus encoding, binary decision, and response production.

As Chiang and Atkinson point out, the sex difference was an unanticipated outcome. It requires replication before conclusions would be warranted. Nonetheless, it is an hypothesis uncovered by careful multivariate exploration of continuous variables. An extreme group design


SATV
800
700
600
500
400
FEMALE
SATV = 508 + 2.15 SLOPE
SATV = 724 − 2.86 SLOPE
MALE
SATM
800
700
600
500
400
SATM = 572 + 1.06 SLOPE
FEMALE
SATM = 790 − 2.51 SLOPE
MALE
FEMALE
MALE
0
20
40
60
80
100
120
SLOPE IN MSEC

leaves out the network of variables in the population from which its groups are extracted. Unless some provisions are made to explore these, important moderating effects are left buried. Without the analysis by sex, the Chiang-Atkinson study would stand as a failure to replicate Hunt's hypothesis.

The Ward Study

We conclude this section with brief mention of a study by Ward (1973). He tested the construct validity of the initial coding component in another way, using a somewhat more elaborate multitrait-multimethod approach. His data bear also on the assumption of independent stages in process models. Individual differences at three processing stages, and their generality across two tasks, were examined. A memory search task following Sternberg (1969) and also a perceptual matching task, where comparison stimuli were present simultaneously, both used drawings of familiar objects, presented on slides under both clear and degraded conditions. Each task allowed measurement of latencies for correct responses to be associated with the effects of three conditions: stimulus degrading, category membership (i.e., match vs. mismatch was correct), and response probability (i.e., match vs. mismatch had high vs. low probability of occurrence). These experimental variables were hypothesized to affect the process stages of stimulus encoding, match-mismatch decision, and response production, respectively. On the memory task, significant main effects occurred for all three experimental variables, and there were no interactions, so the Sternberg hypothesis of independent stages was supported. On the matching task, all main effects were again significant in expected directions, but the interaction of category membership and response probability was also significant. Here then, stimulus encoding could be assumed independent of the next two process stages, but these latter operations seemed not independent of one another. The parameter estimates associated with each stage were similar in the two tasks: stimulus degrading led to an increase in latency of about 200 msec.; a difference of about 70 msec. arose from manipulating stimulus category membership, and change to a lower response probability setting caused an increase in latency of about 55 msec. Correlations between pairs of parameters across the two tasks were 0.75, 0.34, and 0.21, respectively,

implying that individual differences in the encoding stage at least were consistent across tasks. The cross-task correlation for encoding became 0.84 when disattenuated. Several correlations also suggested that the stimulus encoding parameter of one task predicted matching and response production parameters of the other. Thus, while individual differences in the latter parameters seem not to be correlated across tasks, they may represent systematic variance within tasks. Ward's data are partially consistent with our expectations, noted in a previous section, that early stages would be independent, later stages would not, and individual differences would be cumulative across stages.

Future Research

There is other relevant literature, but this is not the place for a comprehensive review. Some brief notes on possible next steps for research in this area will serve as summary.

Hypotheses. Several related hypotheses deserve attention in further research on individual differences in short term visual memory. First is the question of whether individual differences here should be regarded as a function of two (or more) independent parameter differences, or of a single difference in the sequencing of component processing steps. Second, the independence of individual differences in iconic decay, symbolic encoding, single item comparison, and decision and response production must also be checked further. The data so far suggest that the first three steps may be independent and the last two not.

Design of laboratory experiments. A next study in this direction could be conceived as a correlational survey, including measures of all of the above. But a crucial issue concerns the two p-variable vs. one q-variable hypothesis. This appears to require the design of task conditions that can be manipulated to clarify such a distinction, and its inclusion with the above measures in an S-R-R experiment. R. Sternberg's componential analysis may provide a format for the conduct of this work.

The design of such studies will need to include multiple measures of each process construct of interest. And statistical analysis will need to examine possible interactions with sex.

Relevant instructional conditions. The area of short-term visual memory was chosen for demonstration here because data on this were in hand. In general, however, aptitude process analyses should be linked somehow to likely instructional alternatives. Thus the question: Assuming, say, that differences in iconic decay and symbolic coding processes are associated with ability test scores of the PI and V sort, what instructional situations might provide sites where this information is relevant? We might speculate that skilled performance in receiving coded messages visually, in speed reading, and in rapid comprehension of complex figures, pictures, or sequences of these, might depend in part on such individual differences in the short term temporal characteristics of visual processing. Rapid arithmetical computation might also depend on such individual differences, as might various psychomotor skills. Instruction designed to promote learning of these skills could be adapted differently for the High V-Low PI learner and the Low V-High PI learner, by attempting to capitalize on speedy coding or lasting iconic imagery. In general, however, process analysis ought to be guided by an ATI hypothesis from the start, if it is to feed instructional research and development. (See Cronbach & Snow, in press, for related discussion of strategy in ATI research on instruction.)

Capitalization and Compensation: A Challenge

There is no fitting conclusion for the present paper. It has projected one form of suggested further research based on one view of the starting points. There are certainly other views and other starting points. The proof will be in the production of a coherent, useful theory of aptitude based on the orientation provided here, or on one that is demonstrably superior.

In several sections of this paper, various hypotheses have been stated briefly, or implied without explication. There is no sensible alternative at this stage of the work. This is, after all, only a prospectus for further research. But one theme deserves a clearer identification. While introduced late in this paper, it is an overriding theme in all ATI research, and thus a fitting point with which to conclude.

Individual differences among human beings come into play upon situational demand. Individuals seem to meet these demands by capitalizing on their aptitudes, and by compensating for their inaptitudes. Where possible, in effect, they substitute aptitudes they possess for those they lack. In the same sense, situations can be said to capitalize on some individual differences and to compensate for others. The practical problem to which all research on aptitude is ultimately addressed is the design of situations in such a way as to capitalize upon and/or to compensate for the existence of these individual differences. Capitalization and compensation thus seem to be general functions of persons, of situations, and of person-situation interactions. Research on aptitude will need to build a process theory powerful enough to control these functions for practical use.

Allison, R. B. Learning parameters and human abilities. Unpublished report, Educational Testing Service, 1960. (NR 151-113)

Anderson, J. R., & Bower, G. H. Human associative memory. Washington, D. C., Hemisphere, 1974.

Atkinson, R. C., Brelsford, J. W., & Shiffrin, R. M. Multiprocess models for memory with applications to a continuous presentation task. Journal of Mathematical Psychology, 1967, 4, 277-300.

Atkinson, R. C., Holmgren, J. E., & Juola, J. F. Processing time as influenced by the number of elements in a visual display. Perception & Psychophysics, 1969, 6, 321-326.

Averbach, E., & Coriell, A. S. Short-term memory in vision. Bell System Technical Journal, 1961, 40, 309-328.

Berlyne, D. E. Structure and direction in thinking. New York: Wiley, 1965.

Binet, A., & Simon, T. The development of intelligence in children. Translated by E. S. Kit. Baltimore, MD: Williams & Wilkins, 1916.

Bower, G. H. Cognitive psychology: An introduction. In W. K. Estes, (Ed.), Handbook of learning and cognitive processes Vol. I. Hillsdale, N. J.: Erlbaum, 1975.

Breitmeyer, B. G., & Ganz, L. Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. Psychological Review, 1976, 83, 1-36.

Calfee, R. C. Sources of dependency in cognitive processes. In D. Klahr, (Ed.) Cognition and instruction. Hillsdale, N. J.: Erlbaum, 1976

Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.

Carroll, J. B. Psychometric tests as cognitive tasks: A new "structure of intellect". In L. Resnick, (Ed.) The nature of intelligence. Hillsdale, N. J.: Erlbaum, 1976.

Cattell, R. B. Abilities: Their structure, growth, and action. Boston: Houghton-Mifflin, 1971.

Chiang, A., & Atkinson, R. C. Individual differences and interrelationships among a select set of cognitive skills. Memory & Cognition, 1976, 4· 661-672.

Coan, R. W. Facts, factors, and artifacts: The quest for psychological meaning. Psychological Review, 1964, 71, 123-140.

ERIC

Cronbach, L. J. The two disciplines of scientific psychology. American Psychologist, 1957, 12, 671-684.

Cronbach, L. J. Essentials of psychological testing (3rd Ed.). New York: Harper, 1970.

Cronbach, L. J. Test validation. In R. L. Thorndike (Ed.), Educational measurement (2nd Ed.). Washington, D. C.: American Council on Education, 1971.

Cronbach, L. J. Beyond the two disciplines of scientific psychology. American Psychologist, 1975, 30, 116-127.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. The dependability of behavioral measurements: Theory of generalizability for scores and profiles. New York: Wiley, 1972.

Cronbach, L. J., & Snow, R. E. Aptitudes and instructional methods: A handbook for research on interactions. New York: Irvington, 1977.

Duncanson, J. P. Intelligence and the ability to learn. Princeton, N. J.: Educational Testing Service RB 64-29, 1964.

Estes, W. K. Learning theory and mental development. New York: Academic Press, 1970.

Estes, W. K. Learning theory and intelligence. American Psychologist, 1974, 29, 740-749.

Estes, W. K. The state of the field: General problems and issues of theory and metatheory. In W. K. Estes (Ed.), Handbook of learning and cognitive processes, Vol. I. Hillsdale, N. J.: Erlbaum, 1975.

Ferguson, G. A. On learning and human ability. Canadian Journal of Psychology, 1954, 8, 95-112.

Ferguson, G. A. On transfer and the abilities of man. Canadian Journal of Psychology, 1956, 10, 121-131.

Flammer, A. Individuelle unterschiede im lernen. Basel, Germany: Beltz, 1975.

Frederiksen, C. H. Abilities, transfer, and information retrieval in verbal learning. Multivariate Behavioral Research Monographs, No. 2, 1969.

Freeman, F. S. Theory and practice of psychological testing. New York: Holt, 1955.

French, J. W., Ekstrom, R., & Price, L. Kit of reference tests for cognitive factors. Princeton, NJ: Educational Testing Service, 1963.

Gagné, R. M. (Ed.). Learning and individual differences. Columbus, OH: Merrill, 1967.

Gagné, R. M. Learning hierarchies. Educational Psychologist, 1968, 6, 1-9.

Gagné, R. M. The conditions of learning (2nd Ed.). New York: Holt, Rinehart & Winston, 1970.

Garrett, H. E. A developmental theory of intelligence. American Psychologist, 1946, 1, 372-378.

Gaudry, E., & Spielberger, C. D. Anxiety and educational achievement. New York: Wiley, 1971.

Gavurin, E. I. Anagram solving and spatial aptitude. Journal of Psychology, 1967, 65, 65-68.

Gibson, J. J. (Ed.). Motion picture testing and research. Washington, D. C.: U. S. Government Printing Office, 1947.

Glanzer, M. Individual performance, R-R theory, and perception. In R. M. Gagné (Ed.), Learning and Individual Differences. Columbus, OH: Merrill, 1967

Glaser, R. Some implications of previous work on learning and individual differences. In R. M. Gagné (Ed.), Learning and individual differences. Columbus, OH: Merrill, 1967.

Glaser, R. Individuals and learning: The new aptitudes. Educational Researcher, 1972, 1, 5-12.

Glaser, R. Components of a psychology of instruction: Toward a science of design. Review of Educational Research, 1976, 46, 1-24.

Glaser, R., & Resnick, L. Instructional psychology. Annual Review of Psychology, 1972, 23, 207-276.

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

Guinaugh, B. J. An experimental study of basic learning ability and intelligence in low-socio-economic populations. Unpublished doctoral dissertation, Michigan State University, 1969.

Gulliksen, H. O. An analysis of learning data which distinguishes between initial preference and learning ability. Psychometrika, 1942. 7, 171-194.

Gulliksen, H. O, Measurement of learning and mental abilities. Psychometrika. 1961, 26, 93-107.

Guttman, L. The structure of relations among intelligence tests. Proceedings, 1964 Invitational Conference on Testing Problems. Princeton, N. J.: Educational Testing Service, 1965.

Haynes, J. R. Hierarchical analysis of factors in cognition. American Educational Research Journal, 1970, 7, 55-68.

Horn, J. L. Human abilities: A review of research and theory in the early 1970s. Annual Review of Psychology, 1976, 27, 437-485.

Hull, C. L. The concept of the habit-family hierarchy and maze learning. Psychological Review, 1934, 41, 33-52; 134-152.

Humphreys, L. G. The organization of human abilities. American Psychologist, 1962, 17, 475-483

Hunt, D. E. Person-environment interaction. Review of Educational Research, 1975, 45, 209-230.

Hunt, D. E., & Sullivan, E. V. Between psychology and education. Hinsdale, IL: Dryden, 1974.

Hunt, E.B.,& Lansman, M. Cognitive theory applied to individual differences. In W. K. Estes (Ed.), Handbook of learning and cognitive processes Vol.I. Hillsdale, NJ: Erlbaum, 1975.

Hunt, E. B., Frost, N., & Lunneborg, C. E. Individual differences in cognition: A new approach to intelligence. In G. Bower (Ed.), Advances in learning and memory (Vol. 7). New York: Academic Press, 1973.

Hunt, J. McV. Intelligence and experience. New York: Ronald Press, 1961.

Jacobs, P. I., & Vandeventer, M. The learning and transfer of double-classification skills by first graders. Child Development, 1971, 42, 149-159. (a)

Jacobs, P. I., & Vandeventer, M. The learning and transfer of double-classification skills: A replication and extension. Journal of Experimental Child Psychology, 1971, 12, 240-257. (b)

Jensen, A. R. Can we boost intelligence and academic achievement? Harvard Educational Review, 1969, 39, 1-123.

Jensen, A. R. Hierarchical theories of mental ability. In W. B. Dockrell (Ed.), On intelligence. London: Methuen, 1970.

Jensen, A. R. Genetics and education. London: Methuen, 1972.

Kagan, J., & Kogan, N. Individual variation in cognitive processes. In P. H. Mussen (Ed.), Carmichael's manual of child psychology. Vol. 1. New York: Wiley, 1970.

Kahneman, D. Method, findings, and theory in studies of visual masking. Psychological Bulletin, 1968, 70, 404-425.

Klahr, D. (Ed.). Cognition and instruction. Hillsdale, NJ: Erlbaum, 1976.

Koran, M. L., Snow, R. E., & McDonald, F. J. Teacher aptitude and observational learning of a teaching skill. Journal of Educational Psychology, 1971, 62, 219-228.

Maltzman, I. Individual differences in "attention": The orienting reflex. In R. M. Gagné (Ed.), Learning and individual differences. Columbus, OH: Merrill, 1967.

McGeogh, J. A. & Irion, A. L. The psychology of human learning. Toronto: Longmans Green, 1952.

McGuire, W. J. The Yin and Yang of progress in social psychology: Seven koan. Journal of Personality and Social Psychology, 1973, 28, 446-456.

Melton, A. W. Individual differences and theoretical process variables. In R. M. Gagné (Ed.), Learning and individual differences. Columbus, OH: Merrill, 1967.

Merrifield, P. R. Structuring mental acts. In W. B. Dockrell (Ed.). On intelligence. Toronto: Ontario Institute for Studies in Education, 1970.

Mischel, W. Towards a cognitive social learning reconceptualization of personality. Psychological Review, 1973, 80, 252-283.

Neisser, U. Cognitive psychology. New York: Appleton-Century-Crofts, 1967.

Newell, A. You can't play 20 questions with nature and win. In W. G. Chase (Ed.), Visual information processing. New York: Academic Press, 1972.

Newell, A., & Simon, H. A. Human problem solving. Englewood Cliffs, NJ: Prentice-Hall, 1972.

Owens, W. A. Toward one discipline of scientific psychology. American Psychologist, 1968, 23, 782-785.

Reitman, J. S. Computer simulation of an information-processing model of short-term memory. In D. A. Norman (Ed.) Models of human memory. New York: Academic Press, 1970.

Reitman, W. R. Cognition and thought. New York: Wiley, 1965.

Resnick, L. (Ed.). The nature of intelligence. Hillsdale, NJ: Erlbaum, 1976.

Rohwer, W. D., Jr. An introduction to research on individual and developmental differences in learning. In W. K. Estes (Ed.). Handbook of learning and cognitive processes, Vol. 3: Approaches to human learnand motivation. Hillsdale. NJ: Erlbaum, 1975.

Schoenfeld, W. N., & Cummings, W. W. Behavior and perception. In S. Koch (Ed.), Psychology: A study of a science. Vol. 5. New York: McGraw-Hill, 1963.

Seibert, W. F., Reid, J. C., & Snow, R. E. Studies in cine-psychometry II: Continued factoring of audio and visual cognition and memory. Lafayette, IN: Purdue University Audio Visual Center, 1967.

Seibert, W. F., & Snow, R. E. Studies in cine-psychometry I: Preliminary factor analysis of visual cognition and memory. Lafayette, IN: Purdue University Audio Visual Center, 1965.

Simon, H. A. Identifying basic abilities underlying intelligent performance of complex tasks. In L. B. Resnick (Ed.), The nature of human intelligence. Hillsdale, NJ: Erlbaum, 1976.

Snow, R. E. Research on aptitudes: A progress report. Technical Report #1, Aptitude Research Project, School of Education, Stanford University. Stanford, CA: 1976.

Snow, R. E., Marshalek, B., & Lohman, D. Correlation of selected cognitive abilities and cognitive processing parameters: An exploratory study. Technical Report #3, Aptitude Research Project, School of Education, Stanford University. Stanford, CA: 1976.

Snow, R. E., & Seibert, W. F. Exploratory factor analysis in cine-psychometry. Paper presented at the meeting of the American Psychological Association, New York, September 1966.

Spearman, C. The nature of intelligence and the principles of cognition. London: Macmillan, 1923.

Spence, K. W. The nature of theory construction in contemporary psychology. Psychological Review, 1944, 51, 47-68.

Spence, K. W. The relation of learning theory to the technology of education. Harvard Educational Review, 1959, 29. 84-95.

Sperling, G. A model for visual memory tasks. Human factors, 1963, 5, 19-31.

Stake, R. E. Learning parameters, aptitudes, and achievement. Psychometric Monographs, 1961, Whole No. 9.

Sternberg, R. The componential analysis of human abilities: Intelligence, information processing, and analogical reasoning. Unpublished doctoral dissertation, Stanford University, 1975.

Sternberg, S. Memory-scanning: Mental processes revealed by reaction-time experiments. American Scientist, 1969, 57, 421-457.

Stoddard, G. D. The meaning of intelligence. New York: Macmillan, 1943.

Taylor, D. A. Stage analysis of reaction time. Psychological Bulletin, 1976, 83, 161-191.

Terman, L. M. The measurement of intelligence. Boston: Houghton Mifflin, 1916.

Turvey, M. T. On peripheral and central processes in vision: Inferences from an information-processing analysis of masking with patterned stimuli. Psychological Review, 1973, 80, 1-52.

Underwood, B. J. Psychological research. New York: Appleton-Century Crofts, 1957.

Underwood, B. J. Individual differences as a crucible in theory construction. American Psychologist, 1975, 30, 128-140.

Vale, J. R., & Vale, C. A. Individual differences and general laws in psychology: A reconciliation. American Psychologist, 1969, 25, 1093-1108. 1093-1108.

Vernon, P. E. The structure of human abilities. London, Methuen, 1951.

Vernon, P. E. Ability factors and environmental influences. American Psychologist, 1965, 20, 723-733.

Ward, W. C. Individual differences in information processing units. Research Bulletin 73-70. Educational Testing Service, Princeton, NJ: 1973.

Whitcomb, W. A., & Travers, R. M. W. A study of within-test learning function as a determinant of total score. Educational and Psychological Measurement, 1957, 17, 86-97.

Witkin, H. A. The role of cognitive style in academic performance and in teacher-student relations. Princeton, NJ: Educational Testing Service, 1973.

Woodrow, H. The ability to learn. Psychological Review, 1946, 53, 147-158.