DOCUMENT RESUME

ED 198 175 TM 810 143

AUTHOR	Glaser, Robert: Pellegrino, James W.
TITLE	Cognitive Process Analysis of Aptitude: The Nature of Inductive Reasoning Tasks.
INSTITUTION	Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
REPORT NO	LRDC-1979/24
PUB DATE	79
NOTE	28p.: This is the original English version of a paper published in French in the Bulletin de Psychologie, 1979, Volume 32, pp. 603-615.
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Aptitude Tests: Cognitive Processes: *Cognitive Style: *Induction: Intelligence Tests: Learning Theories: *Models: Psychometrics: *Task Analysis

ABSTRACT

This paper provides an overview of the cognitive process analysis of tasks used to measure aptitude and intelligence. As an illustration of this approach, performance on inductive reasoning tasks such as series extrapolation and analogy problems is considered in terms of the factors that contribute to item difficulty and individual differences in measured aptitude. The analysis emphasizes the development of information processing models that specify content and process factors that determine the speed and accuracy of task performance. Theoretical and empirical efforts lead to the postulation of rule complexity and representational variability as two general sources of difficulty in inductive reasoning tasks. These two factors provide a basis for the analysis of individual differences in aptitude test performance and also for instructional experimentation designed to influence reasoning skill. In general, the paper illustrates the implications of applying theories of cognition to the field of psychometrics. (Author)

LEARNING RESEARCH AND DEVELOPMENT CENTER

1979/24

COGNITIVE PROCESS ANALYSIS AND APTITUDE: THE NATURE OF INDUCTIVE REASONING TASKS

ROBERT GLASER AND JAMES W. PELLEGRINO

University of Pittsburgh

COGNITIVE PROCESS ANALYSIS OF APTITUDE: THE NATURE OF INDUCTIVE REASONING TASKS

Robert Glaser and James W. Pellegrino

Learning Research and Development Center

University of Pittsburgh

1979

Preparation of this paper was supported by the Learning Research and Development Center, University of Pittsburgh, which is in part supported by the National Institute of Education, U. S. Department of Health, Education, and Welfare. The opinions expressed do not necessarily reflect the position or policy of NIE, and no official endorsement should be inferred.

This is the original English version of a paper published in French in the Bulletin de Psychologie, 1979, 32, 603-615. The paper is based in part on a paper originally presented at an ONR Research Conference on "Aptitude, Learning, and Instruction: Cognitive Process Analyses," San Diego, March 1978. The most complete description of the research described herein can be found in the paper by J. W. Pellegrino and R. Glaser entitled "Components of Inductive Reasoning." In R. E. Snow, P. A. Federico, and W. E. Montague (Eds.), Aptitude, learning and instruction: Cognitive process analysis of aptitude (Vol. 1). Hillsdale, N. J.: Lawrence Erlbaum Associates, in press. Acknowledgement is due to several former students, namely, Albert Ingram, Joan Heller, and Thomas Holzman for their contributions as cited in the text to the research reported. The authors would like to express their appreciation to Jeffrey Bisanz for his helpful comments on a draft of this manuscript.

Abstract

This paper provides an overview of the cognitive process analysis of tasks used to measure aptitude and intelligence. As an illustration of this approach, performance on inductive reasoning tasks such as series extrapolation and analogy problems is considered in terms of the factors that contribute to item difficulty and individual differences in measured aptitude. The analysis emphasizes the development of information processing models that specify content and process factors that determine the speed and accuracy of task performance. Theoretical and empirical efforts lead to the postulation of _rule complexity_ and _representational variability_ as two general sources of difficulty in inductive reasoning tasks. These two factors provide a basis for the analysis of individual differences in aptitude test performance and also for instructional experimentation designed to influence reasoning skill. In general, the paper illustrates the implications of applying theories of cognition to the field of psychometrics.

COGNITIVE PROCESS ANALYSIS OF APTITUDE: THE NATURE OF INDUCTIVE REASONING TASKS

Robert Glaser and James W. Pellegrino

Learning Research and Development Center
University of Pittsburgh

At the present time, cognitive psychologists are studying individual differences in intelligence and aptitude in terms of the structures and processes hypothesized in the study of cognition and cognitive development. The research carried out along these lines is predicated on the assumption that aptitude tests should be viewed as more than primarily predictors of achievement. Rather, such tests should assist in identifying the processes involved in intellectual competence, and further indicate how these processes can be influenced and utilized to benefit learning. In order to contribute to this purpose, we have undertaken a research program in which we are attempting to identify directly the cognitive processing components of performance on tasks used to assess aptitude. Performance on psychometric test tasks that predict success in education and training becomes the object of theoretical and empirical analyses. The immediate goal is to analyze test tasks, develop process models of task performance, and utilize these models as a basis for individual difference analysis. The distant goal is that, based on this knowledge, conditions designed for learning could be adjusted to these individual characteristics, or instruction could be designed so that it directly or indirectly teaches the processes that facilitate learning. The instructional objective is not to train individuals to score higher on tests of mental ability, but to directly or indirectly influence the cognitive processes that underlie successful performance both on tests of aptitude and in instructional settings.

As a general background for the summary report of our research, we will briefly mention four operating constraints that are important in a task-analytic effort aimed at understanding individual differences in aptitude processes.

A Framework for Task Analysis

First, since there is an extremely large constellation of psychometric tasks, one basic principle for task selection is pervasiveness and robustness. By this, we mean identifying a core set of tasks that frequently occur in widely used tests, and that have demonstrated consistent relationships in factor-analytic studies of certain basic aptitude constructs. A particular task or set of tasks chosen for analysis should have a strong history of reliable association with an aptitude construct that is of reasonable generality, and also has consistent predictive validity with respect to a criterion performance of significant interest.

Second, in the analysis of any particular aptitude construct, it is important to simultaneously consider several tasks or task forms that load on that factor. Analysis should consider the various tasks that intercorrelate and thereby define more completely a substantial set of performances that comprise an aptitude construct. The analysis of multiple tasks should enable us to differentiate general and specific cognitive processes, and thereby help direct us to a level of analysis where research can identify the extent of process trainability and transfer effects.

Third, the analysis of a particular task must also be explicitly concerned with explicating the sources of item difficulty that provide the basis for individual variation in test performance. Test tasks are composed of item sets where the individual items vary considerably in difficulty as a function of ability or developmental level, and an understanding of individual differences in task performance must include a process theory of item difficulty. For this purpose, the processes specified as the components of performance must involve a level of analysis that is sufficient to explain individual item characteristics, individual subject performance, and the interaction of the two.

A fourth constraint arises from the primary goals of a particular research program. In our own work, the goal is to develop a theory of individual differences in aptitude processes that might identify instructionally tractable components of cognitive performance. This constraint is a nontrivial criterion, since it is not unusual for psychologists to engage in detailed task analyses specifying molecular levels of processing that may have little relevance to instructional issues. With this in mind, we propose that the empirical and theoretical results of any particular analysis of the cognitive components of a task be evaluated by asking whether such results suggest testable instructional hypotheses. Such internal tests of the task-analytic effort can be sobering indications that we have yet to achieve a sufficiently useful level of analysis.

The Relevance of Rule Induction

With this general framework of constraints in mind, we will discuss research that our group and others have conducted on a class of tasks that is presumed to assess a psychological capacity for rule induction and is commonly found on tests of aptitude and intelligence. This set of intercorrelated tasks involves several task forms such as classification, series extrapolation, analogy, and matrix tasks. These task forms simultaneously vary along content dimensions that include letters, numbers, words, and geometric figures as shown in Figure 1. Spearman (1923) considered such tasks as measures of *g*, and he viewed them as an index of the capacity to engage in intellectual processes that he referred to as the "education" of relations and correlates. Thurstone (1941) treated these tasks as representative of a primary mental ability called Induction (I), and suggested that rule induction as a second-order factor might be identical with Spearman's *g*. In more recent hierarchical aptitude models, such tasks have been treated as measures of g_f or fluid analytic ability (Brody & Brody, 1976). It seems clear that such rule induction tasks assess basic reasoning abilities that comprise a robust aptitude construct that has relevance for a larger domain of human performance. It has been argued that rule induction processes are similar to those demanded in concept formation, and that they are related to a major form of human problem solving that results in the acquisition of knowledge.

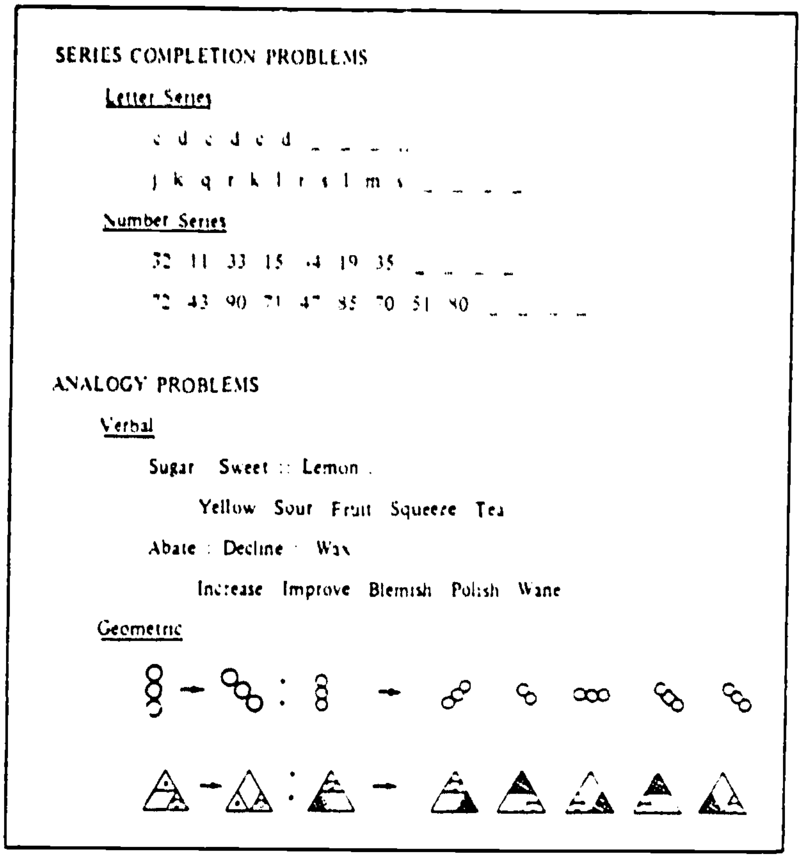

Figure 1. Simple and complex forms of series completion and analogy tasks.

Our tentative view, then, of inductive reasoning tasks of the type found on mental ability tests is that such tasks represent performance samples of the way in which an individual makes use of existing knowledge to solve circumscribed problems where solution depends upon an analysis of the underlying relations (or conceptual similarity) among a set of problem elements. Within important limitations, performance on these tasks has consistently correlated with academic achievement, and individual differences in the capacity to engage in such analyses appear to have direct implications for commonly required classroom learning processes.

We shall now describe a number of research findings on various inductive reasoning test tasks. This research is presented in the form of succinct descriptions and conclusions with little of the experimental detail and caveats which are in available papers (see references cited below).

Series Completion Problems

Series completion items as shown in Figure 1 are found at several developmental levels on many standardized aptitude tests. Such items may be represented as letter series, number series, picture series, or geometric figure series problems. In all cases, the task structure is the same; elements comprising the series are ordered according to some specific interitem relationship, and the individual's task is to extract the basic relationships and generate, predict, or select the next item(s) in the series. The acquisition of serial pattern concepts has an extensive history of psychological investigation. Of particular interest for our purpose here is the work of Simon and Kotovsky (1963; Kotovsky & Simon, 1973) on the analysis of letter series problems of the type developed by Thurstone and Thurstone for their Primary Mental Abilities test battery. Simon and Kotovsky studied adult and adolescent performance in this type of task, and they developed a computer simulation model to represent the component processes necessary for solution.

One important aspect of the analysis of this task is the distinction between the declarative knowledge and the procedural knowledge or processes necessary for the task. The declarative knowledge base for such letter series problems is limited to knowledge of alphabetic order and to relational concepts such as identity (same letter), next (the next letter), and backwards-next (or reverse ordering). Obviously, letter series problems do not involve an extensive declarative knowledge component, and it would not be expected that individual differences would arise from declarative knowledge deficiencies.

Given that the appropriate declarative knowledge is available, the completion of any letter series problem requires a set of basic procedures which are hierarchically organized. In the simulation model, there are two basic routines: a pattern generator and a sequence generator. The first of these routines, pattern generation, can be broken down into three processes: (a) detection of the interletter relations for the given problem elements, (b) use of the relational information to extract the period length of the pattern within the problem, and (c) generation of a pattern description or rule involving both the relations and the periodic structure of the problem. This rule specifying the pattern description serves as input to the sequence generator which applies it to the current state of the problem, and then extrapolates the pattern to generate the additional elements required for problem solution. Differences in item difficulty and potential individual differences in problem solution can result from the application of any or all of these specific processes.

Concerning sources of task difficulty, a number of systematic properties of the individual items determine the difficulty of a problem. One aspect related to the probability of error is the type of relation involved. Identity relations are easier to detect than next relations, which, in turn, are easier than backwards-next relationships. The difference in difficulty between extrapolating identity and next relationships also varies as a function of the position of the relationship within a period.

These sources of error are readily explainable if one considers the requirements of working memory. Identity relationships do not place demands on working memory, whereas successive nonidentity relationships involve accumulating placekeepers. The longer the period length, the greater the memory demands of a problem and the greater the likelihood that working memory limits may be reached. For example, the letter and number series problems in Figure 1 differ in both period length and pattern complexity. The overall pattern description that constitutes a problem is thus related to problem difficulty. The length of the pattern description, which is a function of period length, the types of relationships involved, and the resulting working memory requirements, are highly correlated with problem errors. These errors can arise from performance inadequacies that differ among individuals in relation detection, discovery of the periodic structure, completion of a pattern description, or in the extrapolation process.

Consider now instructional experimentation based upon this model. It has been demonstrated that Simon and Kotovsky's model of serial completion items provides a reasonable account of performance in this task, and their simulation of human protocols provides a partial validation of the model. Given now a concern with the criterion of instructional tractability, another way can be considered in which such models can be validated. The reasoning runs as follows: If the processes embodied in a simulation model are similar to those used by humans, then those processes may be trainable for individuals whose performance represents a low or intermediate level of task competency. Such training should improve performance if the component processes specified and taught are compatible with human cognitive structures. However, if the processes are incompatible with such structures, then there will either be difficulty in training these processes, or they will have no positive and perhaps a negative influence on performance.

In an attempt to provide such an instructional test, a study was conducted that involved direct and independent training in discovery of relations and discovery of periodicity with a sample of children from grades one through six (Holzman, Glaser, & Pellegrino, 1976). Both the training group and a control group were then given a pre- and posttest set of letter series problems identical in rule structure, but initialized at different points in the alphabet. The results showed a significant gain in performance for the training group relative to the control group. In particular, the training group showed a percentage reduction in errors over twice that shown by the control group (32% vs. 13%). Furthermore, in terms of qualitative changes in solution performance, the training group showed significantly greater gains on problems requiring the more difficult patterns, whereas the control group remained the same or reduced errors only on the easier items. Thus, the training appropriately functioned where it was most needed--that is, when individuals encountered more difficult relations and problems. The qualitative difference between control and experimental conditions also suggests that explicit training on the identified component processes may have provided an information management strategy that facilitated pattern description and extrapolation.

Related instructional studies that were conducted on series completion performance showed an interaction between type of training and performance level. Sheer practice on the test items was sufficient to produce significant gains in performance at certain levels of initial competence, but explicit process training was more effective than practice at lower levels of initial ability. This source of interaction needs to be more precisely determined by explicit analysis of the process differences among individuals that define different levels of task competence and intellectual development.

The training study just described represents a very naive form of instructional experimentation. No attempt was made to determine the particular components of performance that were responsible for the different levels of task competence. The process training that was administered was the same for all individuals, and no real effort was made to match training to performance needs. But there is no particular reason why future instructional studies cannot use models of task performance as a basis for both the initial diagnostic assessment of individual cognitive strengths and weaknesses and the subsequent training designed to improve cognitive skills. But so far, we have at least shown that a model of series completion performance can suggest component processes that contribute to item difficulty and that can be translated into a set of procedures for explicit process training. What remains to be shown is whether models of this kind provide useful frameworks for the analysis of individual differences in more generalizable inductive reasoning skill.

We turn now to another task form presumed to assess the capability for Inductive Reasoning.

Analogical Reasoning Problems

Of the many tasks that are assumed to assess this capability, the analogy problem is the most pervasive. Analogy items as shown in Figure 1 have constituted a significant portion of intelligence tests over the entire course of the testing movement. Burt introduced the task in its familiar "A:B::C:D" format in a test published in 1911. Recently, Robert Sternberg (1977a, 1977b) has provided a detailed review and discussion of the importance of analogical reasoning within the field of differential psychology, and the centrality of this type of reasoning with respect to the concept and measurement of intelligence can be found in the writings of individuals such as Spearman (1923) and Raven (1938).

In the past few years, Spearman's theory has been expanded and refined in the more precise, experimentally founded theory presented by Sternberg. He has proposed and tested a theory of analogical reasoning that specifies several processes that are intended to apply across all analogical reasoning tasks. In Sternberg's analyses, emphasis is placed on developing general models of analogy solution and specifying individual differences in terms of latency parameters for the various processes involved. He obtained multiple correlations between .68 and .87 for estimates of process speed and general reasoning scores from a standardized test battery. However, processing latency for the separate components was not uniformly related to general reasoning ability; in some cases individuals with high general reasoning scores were slower on certain component processes. Also, the qualitative characteristics of these processes remain largely unspecified, and we presumably have a poor understanding of how the processes are executed, the information or content that must be processed, and how such information contributes to differences in item difficulty and performance errors.

In our analysis of individual differences in analogy solution, we have attempted to determine the components of task performance that differentiate between levels of aptitude and how these components interact with differences in item structure or content. The task forms studied include geometric and verbal analogies, which we will consider separately.

Geometric analogy task. A starting point for understanding the declarative and procedural knowledge requirements in geometric analogy solution can be found in an artificial intelligence analysis of this task carried out by Evans (1968). He developed a computer program that solved a subset of the geometric analogies that appeared on the American Council for Education examinations.

The principal operations in Evans' program are the following: (a) decomposing the patterns that comprise the terms of the analogy into subpatterns, (b) determining the transformations that relate the subpatterns in the A-B and C-D pairs of terms, and (c) selecting an answer. Decomposition of the A and B terms occurs first. This is accomplished by comparing the figures and determining the common elements or subpatterns. Next, the program matches the subpatterns in the A and B terms and generates a set of transformation rules. The class of transformations recognized by the program includes: removing and adding constituents, rotating, reflecting, and spatial displacement of figures. Following the identification of the constituents of the terms and the transformations relating them, Evans' program decides among the answer alternatives by comparing the A-B transformation with each of the possible C-D options. The D term for which there is the greatest overlap in transformational rules is selected.

The processes represented in Evans' program are compatible with Sternberg's general theory of analogy solution. In all cases, the individual must encode the analogy terms and infer, match, and test the relationships among sets of terms. These processes are influenced by the amount and type of information that must be processed to solve any given item. In geometric analogy items, our studies indicate that item difficulty is a function of both the number of elements and the number of transformations contained in an item, and that solution accuracy and latency depend on the total set of processes required to: (a) decompose complex figures into constituent elements, and (b) identify and order the transformations applied to each element (Mulholland, Pellegrino, & Glaser, 1977). Identifying the transformations applied to individual elements appears to require more processing time and results in more errors than identifying the elements themselves (elements usually being easily perceived geometric figures). Multiple transformations become particularly difficult when applied to a single element, for example, rotating, flipping, and changing the size of a single shape. Such items may be viewed as requiring an individual to operate on a series of partial products, and maintaining and updating the intermediate products place additional demands on processing resources and memory capacity. With certain transformational combinations, the sequence in which the transformations are applied is particularly critical, and this may constitute a further difficulty since there is a need to maintain order information as well.

The work just discussed, together with the geometric analogy experiments described by Sternberg (1977b) and the artificial intelligence analysis of Evans (1968), represent virtually all of the attempts at describing performance in this task. It is obvious that this is only a beginning at understanding the sources of item difficulty and individual differences in performance. The results are encouragingly clear because the items used in experiments are amenable to relatively precise specifications of the information represented in a problem and the rule to be inferred. However, test items do not always have one unique rule or representation. And the solution of such "ambiguous" items depends heavily on the cognitive representation that an individual provides for the elements of an item. For example, in the second geometric analogy item in Figure 1, the transformation between the triangles of the A-B terms was accomplished (represented) by individuals in different ways: (a) by moving the small "X" down to the empty space in the triangle, or (b) by rotating the entire large triangle 120° clockwise and then flipping it over on a vertical axis. Each transformation when applied to the C term dictates the choice of a different answer that is given in the option set. The completeness of the representation, the method employed to achieve the representation and select an answer, and the role of such factors in individual differences remain large question marks.

Further, it must be noted that performance in this task probably involves a considerable amount of procedural knowledge involving strategic flexibility and expertise. The strategic flexibility may be particularly important when it comes to handling complex items. Some items place severe demands on the processing resources and working memory capacity of an individual, and expertise may include the ability to shift processing strategies in order to circumvent some of the working memory problems that arise when using a strategy that is optimal for less difficult items. (For example, an exhaustive initial inference process applied to the A-B pair, in which all the element transformation relations are initially stored in working memory, may be less efficient than a process in which individual element transformation combinations are assessed one by one across the entire problem.) Whether such strategy shifts occur and their relationships to task proficiency remain speculations.

Verbal analogy task. For geometric and alphanumeric series or analogy problems, the declarative knowledge base necessary for solution is relatively finite and easy to identify. Ambiguity may be associated with possibilities of more than one transformational representation, but each is a definite representation that yields a definite solution. In contrast, verbal analogy problems have tremendous variability in the individual elements that may be encountered and in the representation of elements and their relationships, for example, the interpretation of the word "wax" in the second verbal analogy in Figure 1. Furthermore, identifying a particular rule relating two terms does not mean that one can always apply that rule to generate an exact answer. These aspects of verbal analogy problems make it more difficult to specify the processing components and task factors that contribute to differences in item solution and difficulty.

In attempting to analyze the components of performance, it is initially useful to talk about the base or stem of the item separately from the set of alternative completion terms. Analyses of verbal analogy items reveal that the majority of verbal analogies can be classified as representing a limited set of basic types of semantic relations (e.g., Ingram & Pellegrino, 1977; Whitely, 1976). Items containing certain types of relationships such as location or function are verified as true or false faster and have fewer errors than items involving more abstract relationships such as class membership.

However, this classification process does not immediately lend itself to strong predictions about the difficulty of solving an item, since it does not specify the total set of features that must be processed to define the rule for a given item. The semantic characteristics of the item as represented in the stem constitute only part of the semantic task factors contributing to performance. For any item, there also is a set of possible answer alternatives that vary in their semantic appropriateness with respect to matching the semantic features inferred from the stem of the item. These differences between potentially acceptable responses may be considered as representing a "goodness of fit" or semantic distance factor referred to in studies of semantic memory (e.g., Rumelhart & Abrahamson, 1973; Smith, Rips, & Shoben, 1974).

The semantic appropriateness of alternative answers affect performance through the likelihood of their acceptance, and the difference in features associated with a set of alternatives affect the accuracy and speed of any choice among them. The experimental data that we have collected thus far on the effect of these item features on performance reveal that there is rapid and accurate rejection of semantically inappropriate completion terms, and a slower and less likely acceptance of semantically less appropriate completion terms. If errors are made, they seem to be localized in a later stage of processing where additional processing of ambiguous or complex items occurs.

As a general conclusion, performance in verbal analogy tests is consistent with general models of semantic processing. The acceptance or rejection of any given alternative is a function of the congruence between the semantic features that define the A-B and C-D rules. While this is the case for all items, the time and likelihood of accepting the "best" alternative for an item varies considerably across items. And this variability is partially accounted for by the type of relationship involved and the degree of constraint on the set of possible answers for an item (Pellegrino & Ingram, 1977).

The question that now must be considered is whether these characteristics of performance on test tasks differ as a function of aptitude level. Over all items, the high aptitude individuals (as defined by test scores) respond faster than the low group. However, high aptitude individuals tend to spend more time in processing the initial terms of an item. If this response time is considered relative to the total time spent in processing an item, then a pattern emerges in which skilled individuals spend proportionately more time in initial encoding and inference processes and less time in subsequent decision and response processes (see also Sternberg, 1977b).

If it is assumed that the initial encoding and inference processes involve the specification of a rule based upon a set of features relating the A, B, and C terms of an item, then the specificity of that rule must be different in the case of high aptitude and less-skilled individuals. Two findings confirm this notion. First, the amount of time required to specify a precise rule for an item should vary as a function of the number of features or complexity of the rule. This pattern of longer latencies for more complex and difficult items was observed for the high aptitude individuals, with inconsistent latency differences shown by the less-skilled individuals. Second, a more precisely specified rule should allow one to reject false alternatives rather rapidly, and such a result occurred only in the case of high scoring individuals.

In general, the data suggest that individuals with high aptitude test scores specify more precisely the set of semantic features representing the interrelationships among the individual components of the item, and that this difference in the quality of encoding gives rise to different latency and error patterns. Protocol analyses of individual subjects to further investigate this issue have highlighted a number of performance characteristics that differentiated between items that individuals find to be of high and low difficulty (Heller & Pellegrino, 1978). Items that are relatively easy lend themselves to a solution process in which the relationship is readily specifiable, and a potential completion term for the item is easily generated. Thus, the process of solution follows a generate and test model in which the processing of the alternatives involves a simple search for the hypothesized answer. This can be called a "working-forward" strategy and is illustrated in the protocol shown in Table 1 for the item Tea:Coffee::Bread:____________. In contrast, difficult items for individuals are ones in which the relationship is not well specified, and there is difficulty in generating a potential answer. In this case, solution is guided by the set of alternatives and the relationship is defined by working backward from these alternatives. The protocol of such a solution process is shown in Table 2 for the item Subject:Citizen::King:____________. Thus, items appear to vary on a continuum from a hypothesis generation and test process for easy items where there is little change in the level of feature analysis, through to a process for difficult items that is largely driven by the alternative set with constant redefinition of the possible relevant features of the problem. Thus, depending upon the difficulty of an item for an individual, verbal analogy solution varies from a simple, sequential execution of component processes to complex combinations of working forward and working backward strategies with increasing reliance on the option set. These latter cases either involve ambiguous relationships or the

presence of two or more options at similar degrees of analogical appropriateness, and conceptualization of the relationship often is constructed through a gradual, iterative process.

Table 1

Solution Protocol for "Simple" Analogy

Analogy Elements Presented		Solver's Response
TEA: COFFEE:: BREAD:		Tea is to coffee as bread is to . . . rolls because tea and coffee, they're both drinks, and they're about the same thing, just two different names for two different drinks, and a bread and a roll would be about the same--two different names for the same thing.
MILK	(Reject)	That doesn't fit, it's a drink.
BUTTER	(Reject)	Butter is something you put on bread, that doesn't fit.
ROLLS	(Accept)	That's good.
JAM	(Reject)	It's like butter, something you put on bread. It wouldn't fit because you don't put coffee on tea or in tea.

Table 2

Solution Protocol for "Complex" Analogy

Analogy Elements Presented		Solver's Response
SUBJECT: CITIZEN:: KING:		Subject . . . king . . . If you're subject to something . . . no, that wouldn't have anything to do with it. I don't know about this one, what could fit in there.
	Researcher:	Do you see any relationships between anything?
		Well, the king would sort of subject the people to stuff so maybe it would be something like the people who are under him.
RULE	(Accept)	Citizens are subjects and a king rules--that could be true. I sort of think of it in a way like the citizens are people that are subject to the rule and the king is the one that does the ruling.
KINGDOM	(Reject)	A king would rule over the kingdom--but the citizens are subject--ooou! I don't know, but I think citizens are subjected to the ruler.
PRESIDENT	(Accept)	The king and the president would be the ones that rule. Subjects and citizens would be the ones that they rule, the people that they rule.
KNIGHT	(Reject)	I don't see what knight would have to do with it.
RULE	(Reject)	I'd say that's false because rule is what the king would do. But citizen and subject, it doesn't say what they do, they're sort of like the same thing.
PRESIDENT	(Accept)	A king's like a president, a subject's like a citizen.

Implications

Individual differences and item difficulty. At the beginning of these remarks, certain criteria were stated that are useful for evaluating the results of an analysis of tasks representing a major aptitude construct.

One criterion is whether the results of the analysis move us closer to an analytic scheme that can be used to diagnose performance deficiencies, and a second is whether the sources of individual differences suggest testable instructional hypotheses. As is obvious, this paper has been primarily devoted to theoretical and empirical analyses of the task and item properties that influence performance on aptitude test tasks, and we are only beginning to understand the sources of individual differences. Thus, at this point, we can only speculate about the instructional implications of individual differences in inductive reasoning skills. However, the work so far clearly provides a basis for individual differences analyses, particularly with respect to the sources of task difficulty. There appear to be two major factors that contribute to task difficulty; these include: (a) the complexity of the rule to be inferred, and (b) the variability or initial ambiguity in the possible rules that may be inferred.

In all of the tasks that employ nonverbal stimuli, rule complexity is reflected directly by the number of different operators that must be represented in working memory. The greater the number of transformations that comprise a rule to be inferred and applied, the larger the increase in both solution latency and error probability. This is due to difficulties in assembling and maintaining a complete description of the element-operator combinations. Such descriptions may exceed memory capacity, and only partial representations of the rule to be induced and applied may be established.

In verbal items, rule complexity is more difficult to measure since it is not directly manifested by any overt problem features. Also, verbal items seem to place somewhat different demands on working memory, particularly with respect to maintaining information about past and current hypotheses about the relevant semantic features that comprise an item's rule. The changing nature of the semantic feature set that occurs as new item terms are encountered is related to what can be called representational variability. In verbal analogies, item difficulty is heavily influenced by variability resulting from the conceptual richness and abstractness of the individual terms and relations. Verbal concepts activate extensive semantic feature sets, and varying subsets of these features may be included as part of the problem representation. The ability to entertain different problem representations and modify them during the course of solution represents a significant aspect of performance. This is also true for nonverbal items that often have two or more possible representations for the rule governing the problem.

The factors of rule complexity and representational variability provide a scheme within which to consider sources of individual differences. The limited individual difference data that we have discussed focus primarily on the representational variability factor as manifested in verbal analogies. The data indicate that skill differences in an undergraduate population are associated with: (a) processes of establishing a reasonably well-defined problem representation, (b) the subsequent utilization of that representation as a basis for selecting among alternatives, and (c) modifying the representation as necessary. The time spent establishing an initial representation (or representations) may differ as a function of aptitude level, but latency differences may be less important than the particular representation(s) achieved. Indeed, there is evidence in our data and others' that high aptitude individuals, who presumably have more elaborate semantic memory structures, may encode more item features and take more time in this aspect of processing, but with subsequent facilitation in the speed and accuracy of selecting among alternatives.

Instruction. The different task analyses that we and others have carried out suggest three potential areas for instructional research. But, before making these suggestions, we must remind ourselves about what the goal of our enterprise is. It is certainly not to produce merely high scores on aptitude tests. (This can be attempted by such means as buying a practice booklet for the Miller Analogies Test that provides drill on the classes of relations that appear on that test.) Rather, the eventual goal is to increase individuals' proficiency in inductive reasoning in the context of subject matter knowledge in order to improve learning and attained competence. Essentially, we are interested in a fourth R that should pervade the other three--Reading, Writing, (A)rithmetic, and Reasoning.

With this in mind, the first general instructional research suggestion points to the fact that there is a declarative, content knowledge base that must be available to perform any reasoning task. Declarative knowledge for rule induction tasks involves the basic elements and possible transformations for the particular content area. This is a large, relatively unbounded set of knowledge in the case of verbal reasoning, and thus, it poses a nontrivial instructional problem. Nevertheless, it might be possible to teach the necessary declarative knowledge for a restricted content domain, and then use this to permit the exercise of inductive reasoning processes that involve a search for relations among relevant elements of knowledge that influence further learning.

The second target for instruction would involve factors associated with the rule complexity dimension. This might involve two different types of instruction, one being instruction in processes associated with storing, retrieving, and manipulating information in working memory. Whether such instruction is feasible is unclear, but work is proceeding on understanding these memorial processes. The second potential form of instruction would involve executive strategies for organizing, controlling, and monitoring the analysis of problem features during the course of solution. Such procedures may substantially reduce memory load problems, and they can be viewed as forms of procedural knowledge. Such procedural knowledge may be instructable, and minimal evidence for this is that partial procedures such as how to discover periodicity in series completion problems can be readily taught to children, and such knowledge affects performance within the task.

The third target for instruction involves factors associated with the representational variability dimension. In this case, the form of instruction may be linked to knowledge about general aspects of problem solving such as defining the relevant problem space for a task and then using information within the problem space to help restructure the problem.

The extent to which and how such problem-solving skills can be operationalized and taught is also unclear at present. But our guess is that knowledge of instructional procedures for developing these skills can be obtained from study of the carefully sequenced problem sets that an expert tutor presents to students.

Finally, it can be concluded that systematic task analytic efforts of the cognitive processes involved in sets of tasks such as we have described should help make it possible to understand the link between rule induction skill and instructional situations that influence this skill. In this way, a contribution might be made toward generally improving the success in learning that aptitude tests attempt to predict. Consider the possible relevance and generality of processes that have been identified in the small set of studies of inductive reasoning tasks. The rule induction processes involved in analogical reasoning and series completion appear to be similar to many forms of problem solving and concept formation. The essence of this similarity is the ability to search for relations among elements resulting in new interconnections between concepts stored in memory. Consistent with this contention, it has been argued that one of the learner's essential roles in classroom learning is to recognize the structural form or pattern of the facts conveyed by instruction and to detect relations between this newly communicated material and the material already existing in a semantic network in memory (Norman, Gentner, & Stevens, 1976).

When we are able to specify intellectual abilities in terms of psychological processes, then we have information that enables us to do more than predict performance on a criterion task. We then have information that provides a basis for doing something about performance either by engaging in specific process training designed to improve performance or by changing the learning situation to make the attainment of criterion performance more likely. An increasing number of studies are being carried out to determine the direct instructability of specific processes that underlie intelligence and aptitude. Much of this work is being done in the context of research on mental retardation (Belmont & Butterfield, 1977; Brown, 1974; Campione & Brown, in press). However, testing the limits of such training on a wider spectrum of abilities has not yet begun. Experimental studies investigating the possibility of changing the learning situation to adapt to individual differences have been largely represented by studies of aptitude-treatment interaction. But aptitudes and treatments have been rarely analyzed in terms of the relationships of similar underlying performance processes, and this probably accounts for the many negative findings in this area (e.g., Cronbach & Snow, 1977). Success might be expected when we have a more refined process analysis that relates individual capabilities and learning requirements.

The potential benefits that can be derived from an understanding of the cognitive components of individual differences are consistent with the nature and purposes of education. It is no longer possible to consider testing only as a means of determining which individuals are already adapted to or who have the potential for adapting to mainstream educational practice. Society's current goal is to reverse this sequence of adaptation; rather than requiring individuals to adapt to means of instruction, the desired objective is to adapt the conditions of instruction to individuals to maximize their potential for success (Glaser, 1977).

This objective can be realized if learning can be designed to take account of an individual's profile of cognitive skills. If we analyze the performance requirements of various school activities and then analyze the skills that individuals bring to these task environments, we should be able to match the two. Learning could then be assisted in two ways. First, the cognitive skills that individuals bring to schooling could be matched to various instructional environments that utilize these skills. Second, the cognitive skills of an individual could be improved to meet the demands of the instructional environments that are available. Providing for both the development of cognitive skills and accommodating to different cognitive capabilities offers maximum adaptability for enhancing the likelihood of effective education.

References

Belmont, J. M., & Butterfield, E. C. The instructional approach to developmental cognitive research. In R. V. Kail, Jr., & J. Hagen (Eds.), Perspectives on the development of memory and cognition. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977.

Brody, E. B., & Brody, N. Intelligence: Nature, determinants, and consequences. New York: Academic Press, 1976.

Brown, A. L. The role of strategic behavior in retardate memory. In N. R. Ellis (Ed.), International review of research in mental retardation (Vol. 7). New York: Academic Press, 1974.

Campione, J. C., & Brown, A. L. Toward a theory of intelligence: Contributions from research with retarded children. Intelligence, in press.

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

Evans, T. G. Program for the solution of a class of geometric-analogy intelligence-test questions. In M. Minsky (Ed.), Semantic information processing. Cambridge, Mass.: MIT Press, 1968.

Glaser, R. Adaptive education: Individual diversity and learning. New York: Holt, Rinehart, & Winston, 1977.

Heller, J. I., & Pellegrino, J. W. Cognitive processes and source of item difficulty in the solution of verbal analogies. Paper presented at the meeting of the American Educational Research Association, Toronto, March 1978.

Holzman, T. G., Glaser, R., & Pellegrino, J. W. Process training derived from a computer simulation theory. Memory & Cognition, 1976, 4, 349-356.

Ingram, A. L., & Pellegrino, J. W. Response generation norms for verbal analogies. Pittsburgh: University of Pittsburgh, Learning Research and Development Center, 1977.

Kotovsky, K., & Simon, H. A. Empirical tests of a theory of human acquisition of concepts for sequential patterns. Cognitive Psychology, 1973, 4, 399-424.

Mulholland, T., Pellegrino, J. W., & Glaser, R. Components of geometric analogy solution. Paper presented at the meeting of the Psychonomic Society, Washington, D. C., November 1977.

Norman, D. A., Gentner, D. R., & Stevens, A. L. Comments on learning schemata and memory. In D. Klahr (Ed.), Cognition and instruction. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1976.

Pellegrino, J. W., & Ingram, A. L. Components of verbal analogy solution. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, May 1977.

Raven, J. C. Progressive matrices: A perceptual test of intelligence, 1938, individual form. London: Lewis, 1938.

Rumelhart, D. E., & Abrahamson, A. A. A model for analogical reasoning. Cognitive Psychology, 1973, 5, 1-28.

Simon, H. A., & Kotovsky, K. Human acquisition of concepts for sequential patterns. Psychological Review, 1963, 70, 534-546.

Smith, E. E., Rips, L. J., & Shoben, E. J. Semantic memory and psychological semantics. In G. H. Bower (Ed.), The psychology of learning and motivation (Vol. 8). New York: Academic Press, 1974.

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

Sternberg, R. J. Component processes in analogical reasoning. Psychological Review, 1977, 84, 353-378. (a)

Sternberg, R. J. Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977. (b)

Thurstone, L. L., & Thurstone, T. G. Factoral studies of intelligence. Chicago: University of Chicago Press, 1941.

Whitely, S. E. Solving verbal analogies: Some cognitive components of intelligence test items. Journal of Educational Psychology, 1976, 68, 234-242.