Information-Processing on Intelligence Test Items: Some Response Components

— Source link ↗

Susan E. Whitely

Institutions: University of Kansas

Related papers:

- Solving verbal analogies: Some cognitive components of intelligence test items.

- Component Processes in Analogical Reasoning.

- Intelligence, information processing, and analogical reasoning : the componential analysis of human abilities

- Psychometric Tests as Cognitive Tasks: A New 'Structure of Intellect'.

- Modeling Aptitude Test Validity from Cognitive Components.

Information-Processing on Intelligence Test Items: Some Response Components

Susan E. Whitely
University of Kansas

The reorientation of experimental psychology from studying performance to studying cognitive processes has created a new potential for understanding ability tests in terms of the nature of the cognitive events which contribute to individual differences in solving the test items. The results from the present study suggest the feasibility of explaining individual differences in performance on a prototypic intelligence test item—verbal analogies—from the success and efficiency of processing information on hypothesized component events. The data revealed that at least three types of processing events are needed to describe individual differences in the component task durations, but that probably only one factor is needed to describe accuracy in completing the components. More critically, both the accuracy and duration of the component tasks were significantly related to solving psychometric analogies. The results are discussed with respect to the nature of successful performance on analogy test items and the need for more complex models to fully account for individual differences in test performance.

The nature of the capabilities measured by ability tests has perplexed psychometric theorists for several decades. The split between substantive theory and measurement began with Binet, who used empirical criteria to distinguish between practically important groups of people—rather than theory—to select intelligence test items. Certainly Binet's decision was wise in terms of providing a foundation for measures of intelligence which have had enormous social impact in a variety of applied settings. At the same time, however, decades of research have not led to a satisfactory theoretical understanding of ability tests. As Messick (1972) pointed out, most of the psychometric theories and supporting test correlation data have really not sought to understand intelligence test performance in terms of more basic theoretical entities. That is, the major concern has been trait organization rather than developing models of the psychological processes which underlie intelligence test performance.

In contrast, experimental psychology has been devoted to developing models for understanding intelligent task performance, but not to measuring individual differences. Instead, general theories have been formed to explain intelligent behavior, and then tasks are devised to test the theory.

The different approaches of experimental psychology and psychometrics have resulted in practically important measures of intelligence which are not understood, in addition to cognitive theories which have no practical implications for measuring individual differences.

One attempt to bridge the gap between psychometric instruments and experimental

theories has consisted of correlating ability test scores with the theoretical variables measured from experimental tasks. The positive correlations which were obtained in a series of studies in the 1950s and 1960s (i.e., Allison, 1960; Dunham, Guilford & Hoepfner, 1968) are somewhat encouraging since they suggest that experimentally-based measures and psychometric instruments may inadvertently share some common concerns. However, these studies have not contributed substantially to a theoretical understanding of intelligence tests. Not only were the correlations fairly low, but the task measures themselves were not understood in terms of the components involved in performance.

Recently, however, the turn of experimental psychology from the study of learning and performance to the examination of the nature of the covert cognitive events involved in task completion has created a new potential for understanding ability tests. By observing the effects of type and amount of information input on task reaction times and responses, experimental psychology has developed models of complex cognitive processes. Carroll (1974) suggested that a comparable approach to studying test items may not only link intelligence tests to cognitive theory, but may also explain the trait organization findings. Carroll hypothesizes that correlations between items or tests arise when components of information processing, such as common memory stores or common cognitive operations required to answer the items, are shared. Accordingly, research which examines the information-processing components for different types of aptitude test items could potentially predict and explain their intercorrelations. Carroll believes that studying aptitude test items in this manner will provide a new basis for structure of intellect models in psychometrics.

A recent series of studies (Hunt, Frost, & Lunneborg, 1973; Hunt, Lunneborg, & Lewis, 1975; Lunneborg, 1974, 1977) provides the most extensive data available on some information-processing variables which are related to intelligence test performance. These studies found that scores on a comprehensive verbal intelligence test were significantly related to many as-

pects of information processing, including speed in converting sensory data to conceptual meaning, speed in scanning data in short-term memory, retention of order information and resistance of memory information to interfering data.

However, with respect to identifying the components of information processing involved in solving intelligence test items and in line with Carroll's (1974) suggestion, the Hunt et al. (1973, 1975) and Lunneborg (1974) studies are only suggestive for two reasons. First, these studies used composite intelligence test scores which were based on several distinct types of items. Thus, processes which are correlated with the total score may not be reflected in each type of item. Second, the information-processing tasks in these studies may not have measured those processes that are directly involved in intelligence tests. The laboratory tasks were selected to reflect a variety of contemporary concerns in experimental research rather than to represent component events which may be inherent in solving items. Many tasks, for instance, even used stimulus materials which differed in semantic content from the items on the verbal intelligence test (i.e., some tasks used numbers or symbols). Unfortunately, a change of stimulus content often results in the need for a rather different account of information processing in a task. Thus, the positive correlations found in these studies may not be readily explained as resulting from shared components of information processing.

One method of obtaining more direct data on the processes involved in solving intelligence test items is to study a single type of item. Of the many types of items which appear on popular intelligence tests, verbal analogies are particularly well suited for experimental study for several reasons. First, verbal analogies show the highest saturation on the general factor which is common among intelligence test items. Second, verbal analogies were given a central role in Spearman's (1923) psychometrically oriented theory of intelligence and cognition. Third, each verbal analogy consists of a set of single-word

stimuli, making information-processing research more easily applied to analogies than to more complex comprehension items.

A second method of obtaining more direct data on processing is to use laboratory tasks which represent some component events in solving the test item. Under this approach, the test item pool would serve as the source of word stimuli for the tasks, and the tasks would vary in the component phases of item solving that the subject performs. Two separate dependent variables are of interest for component analysis: (1) accuracy—to measure the probability that the state is successfully completed and (2) response time—to measure the duration of the covert mental events required to complete the stage successfully. Both variables are important, since not only are their intercorrelations rather low, especially for test items (Davidson & Carroll, 1945), but their implications for individual differences are not the same. Although accuracy measures the object difficulty of a component for an individual, response time is more reflective of mental processing.

The goal of the present research is to study general intelligence test items as experimental tasks to improve our understanding of the nature of the component processes that contribute to individual differences in test performance. Similar to the Hunt et al. (1973, 1975) and Lunneborg (1975) studies, the current research correlates laboratory task performance with test scores. However, this study is more directed toward identifying processes involved in solving test items than previous studies on information-processing and intelligence tests because a single type of item, the verbal analogy, is used to measure intelligence, and the laboratory tasks are designed to represent potential component phases of solving the test items.

Method

Design

The basic design of this study was to construct several laboratory tasks representing a series of information-processing stages on analogies, and then to examine the relationship of the stages to individual differences on analogy ability tests through regression analysis. The dependent variable was performance on psychometric analogies, and both number correct and average response latency were used as separate measures of individual differences. The independent variable was performance on the various stages, which was measured by response accuracy and response latency on the hypothesized component tasks. The relative importance of the stage components in explaining the two dependent measures of analogy item performance—response latency and number correct—was determined from predictor contributions in regression analysis. However, prior to the regression analysis, the component tasks were organized into dimensions which defined independent aspects of individual differences through separate factor analyses of the task response times and the task accuracies.

Subjects

The sample consisted of 71 high school seniors, randomly selected from a class list. The school was located in a central area of a large Midwestern city, and the enrollment was approximately 50 percent minority race students, largely Black. In the sample, 37 subjects were minority race students and approximately 52 percent of the sample was female, with the sex distribution approximately equal within race. Lorge-Thorndike Intelligence Test scores, Differential Aptitude Test (DAT) scores and grades were available from the school files.

Equipment

The equipment consisted of a carousal slide projector and a LaFayette Instrument Company Model 5710 timer and automatic printer. The experimenter response button was relayed to reset the timer, change the slide and start a new time interval. The subject response button was

Figure 1
Example of Analogy Item

(STEM)	Race: Man:: Strain:_____
(ALTERNATIVES)	1) Science 2) Bacteria 3) Fatigue 4) Soup 5) Tension

relayed to stop the timer and print the response time for the slide.

Materials

The materials included an analogy test and eight experimental tasks. The *Analogy Test* consisted of 27 analogies selected from an item bank which had been developed and standardized in previous work (Tinsley & Dawis, 1975; Whitely & Dawis, 1974). Analogies in the test were similar to those found on popular intelligence tests such as the Scholastic Aptitude Test or military intelligence tests, except that the vocabulary did not exceed sixth grade level in order to assure appropriateness for central city students. Figure 1 displays the general format of the analogy items.

For the laboratory tasks, it was hypothesized that the component stages could be organized into the following three areas: (*1*) memory processes, i.e., stages which include recoding (converting) the printed analogy stimulus into an internal representation in short-term memory and retrieving previously acquired information about the stimuli from long-term memory; (*2*) control strategies, a series of higher order activities, e.g., evaluating the quality and relevance of information, comparing alternative solutions and determining if further information is needed; and (*3*) response implementation, involving stages which include matching the analogy solution with the alternatives and recording the response. Eight tasks involving ten separate timed measures were constructed to tap these components. With the exception of one task *(Choice Reaction Time)*, all task stimuli were obtained from analogies which were part of the same item bank as the *Analogy Test*. In no case, however, was the same item used in more than one task.

Memory processes. The first major area, memory processes, was represented by two tasks. One of these tasks—*Short-Term Retention*—was used to measure the memory encoding stage in which the individuals input a stimulus into short-term memory so that they may perform further operations. Thus, the short-term retention task was constructed to give a measure of processing duration for encoding the printed analogy stem. In this task, subjects viewed an analogy stem, such as the following:

Deep: Cheap:: Shallow:_____

and they were given instructions to "study the analogy stem so that you can remember it." The task required subjects to reproduce verbally the analogy stem to the experimenter after a five-second study-recall interval. Both recall error and study times were recorded on the ten task items.

The other memory task—*Relationship Education*—was constructed to measure speed in retrieving information about word relationships from long-term memory. This task was named after Spearman's (1923) theoretical description of relationship retrieval as one of three components of intelligent behavior which is prototypically represented by the analogy item. In the task, subjects were required to find a relationship between a pair of words, such as the following:

Deep: Cheap

Subjects were instructed to push the response time button as soon as they thought of a relationship between the words, and then to write a short sentence describing the relationship. Latencies were recorded for each of the ten task items.

Control strategies. Control strategies were represented by two tasks. The first task—*Rela-*

Figure 2
Example of Relationship Evaluation Task

Deep: Cheap:: Shallow: <u>Bargain</u>

1	2	3	4	5	6
Terrible	Poor	Questionable	Adequate	Good	Excellent

tionship Evaluation—required subjects to rate the quality of an analogy on a six-point numerical rating scale. Figure 2 is an example of the task. Subjects were instructed to rate "how well the underlined word completes the analogy."

The item pool data were used to select stem completions in this task. For five analogies, the alternative was the correct answer. For five other analogies, the alternative was a popular distractor, to which more than 35% of the wrong responses were given ($\bar{p} = 54.4\%$), while the remaining five analogies were completed with an unpopular distractor to which less than 30% of the wrong responses were given ($\bar{p} = 26.1\%$). The items in each set were balanced for difficulty with correct items, popular distractor items and unpopular distractor items having difficulties of 55.18, 56.73, and 59.68 percent, respectively.

Although this task obviously involves retrieving information from memory, deciding which word pairs to compare and judging the preciseness of the match between word pairs requires some control processes. Thus, individual differences in response times should reflect control processes as well as memory processes. Study times and ratings were recorded for the fifteen task items.

The second task constructed to measure control processes—*Relationship Choice*—was a comparison task which required subjects to select the best of two completions of the same analogy stem. For example:

   a) Deep: Cheap:: Shallow : <u>Bargain</u>
             OR
   b) Deep: Cheap:: Shallow : <u>Costly</u>

One of the analogies was completed with a correct alternative, while the other was incorrect. For five items, the incorrect alternative was a popular distractor, as defined above ($\bar{p} =$

53.7%), while the other five items had unpopular distractors ($\bar{p} = 27.2\%$). Item difficulties were 55.38 and 57.38% for items with popular and unpopular distractors respectively.

As with *Relationship Evaluation*, memory processes are involved in task completion, but the comparison process is an additional component. Response times and errors were recorded on the ten task items.

Response implementation. In the preceding tasks—*Short-term Retention, Relationship Eduction, Relationship Evaluation* and *Relationship Choice*—the response times did not include the recording of the response. The subject pushed the time switch prior to indicating his/her answer to the task so that the time measures were not confounded with the varied response requirements of the tasks. However, response implementation was hypothesized to represent a major area of the analogy-solving process, so a task was constructed to represent the response process required on standard analogy items. In this task, subjects were shown the answer to an analogy and the five alternatives, as follows:

    Costly
        1) Costly
        2) Wide
        3) Steep
        4) Plenty
        5) Bargain

The subject was instructed to "find the alternative that matches the single word as quickly as you can. When you find it press the response button . . . then circle the number of the correct alternative on the answer sheet. When you finish, press the response button again." Two times were obtained from the two stages of this task, *Response Decision* and *Response Execution*, respectively, for ten task items.

An additional response task, *Choice Reaction Time,* was constructed to represent a baseline response decision task which involves little cognitive mediation. Subjects viewed either a series of +'s or 0's and were instructed to "press the response button as fast as you can when the slide with the +'s appears." Reaction times and errors were recorded for the set of ten task items with +'s.

A last task consisted of a standard analogy item divided into two stages. This task was not specifically designed to represent any of the major components, but was included to examine the effects of presentation strategy on the nature of processing. The subject was first shown the analogy stem, as in the following example,

Deep: Cheap:: Shallow:_____

and instructed to "study it until you understand what kind of alternative will complete the analogy." Then the subject was shown the complete analogy, with alternatives, as follows,

Deep: Cheap:: Shallow:_____
    1) Costly
    2) Wide
    3) Steep
    4) Plenty
    5) Bargain

and asked to "find the alternative that best completes the analogy." Response times were obtained from both stages, and were labeled *Relationship Study* and *Analogy Completion After Study,* respectively. The task consisted of ten items.

Procedure

Each subject was tested individually by an experimenter in a small, quiet room located in the school building. First, the subject was given some general instructions about the equipment. Subjects were told that their response times and response accuracies would be recorded, and that task items should be completed as quickly as possible, while still making the correct response

to the item. Then, the subject was administered the 27-item *Analogy Test* according to standard multiple-choice instructions. The items were shown on slides and response times were recorded. Next, the subjects were administered the eight laboratory tasks with the ten timed measures. The tasks were presented in the following order: *Choice Reaction Time, Response Decision, Response Execution, Short-term Retention, Analogy Study, Analogy Completion After Study, Relationship Eduction, Relationship Evaluation* and *Relationship Choice.* Prior to each task item set, subjects were given examples in the instructions and then given one or two practice items.

Scoring

Both number of correct responses and average response times (accurate to .01 second) were recorded for the standard analogy items and the component tasks. In accordance with standard procedures in response time experiments, an average time for each component task was scored using only the items which the subject answered correctly. Traditionally in information processing, a stage cannot be assumed to occur unless the correct answer is given. For the standard multiple-choice items on the *Analogy Test,* which are the complete information-processing events to be explained by the component tasks, a different procedure was used. Latencies were averaged *separately* for items answered correctly and incorrectly to examine the relative importance of the various stages of information processing in these two different outcomes.

The *Relationship Evaluation* task required scoring the item evaluation ratings as correct or incorrect. Ratings were scored as "correct" if an invalid analogy was rated 1, 2, or 3 or a valid analogy was rated 4, 5, or 6, as determined by whether or not the fourth term in the analogy was the keyed alternative in the item bank.

Missing data were handled by computing the statistics on all available scores rather than eliminating entire protocols. An average of four

Table 1
Means, Standard Deviations and Intercorrelations Among
Background Data and Analogy Test Variables

Variable	$\bar{X}$	SD	Correlations				
			2	3	4	5	6
1. Lorge-Thorndike	99.50	12.95	.84**	.52**	.64**	-.36**	-.15
2. DAT	34.00	14.88		.47**	.70**	-.28*	.06
3. GPA	2.66	.77			.45**	-.17	-.04
4. Analogies, Number Correct	8.10	4.46				-.06	.23*
5. Analogies Latency (Correct)	12.92	4.90					.82**
6. Analogies Latency (Incorrect)	13.07	4.90					

*
Statistically significant at $p \leq .05$
**
Statistically significant at $p \leq .01$

Table 2
Means, Standard Deviations and
Factor Loadings for Task Response Times

Variables	$\bar{X}$	SD	Factor		
			I	II	III
Choice Reaction Time	.72	.36	-.11	.06	.13
Response Decision	2.71	1.71	.05	.76	-.01
Response Execution	2.42	1.31	.14	.98	.06
Short-term Retention	3.49	1.66	.67	.19	.14
Analogy Completion After Study	6.73	3.71	.45	.09	.09
Relationship Study	4.98	2.02	.13	-.04	.73
Relationship Eduction	6.14	3.83	.91	-.06	.21
Relationship Evaluation	8.72	3.66	.56	.01	.73
Relationship Choice	9.22	5.06	.46	.04	.59
Proportion of Common Variance			.40	.31	.29

percent of the data was missing for the tasks, due to equipment failures or recording errors.

Results

Table 1 presents data on some background variables and performance on the multiple-choice *Analogy Test* items. The sample ability level was average on the Lorge-Thorndike Intelligence Test, but below average on the DAT. The average *Analogy Test* score was also somewhat below the standardization sample, with an average item difficulty of 33.75% in the current sample. The latency in solving the *Analogy Test* items was fairly short, less than fourteen seconds for both correct and incorrect items. Correlations with the three analogy variables are also presented in Table 1. The analogy-solving latencies were highly correlated between correct and incorrect items. However, latencies for correct items did not correlate significantly with number of analogies solved (i.e., the *Analogy Test* score), while latencies for incorrect items showed a marginally significant correlation ($p < .05$). The number of analogies solved correlated highly with the general intelligence tests (as expected) and was a moderately good predictor of GPA. For the analogy solving latencies, only the time for the correct items correlated significantly with the Lorge-Thorndike and DAT.

Table 2 presents data on the latencies for the component tasks. The shortest average latency (.72) was obtained for *Choice Reaction Time,* while the longest was found for *Relationship Choice.* A factor analysis was performed to organize the component task latencies into independent dimensions of individual differences. Principal factor analysis, using squared multiple correlations for the initial communality estimates, was applied to the task response-time correlations. Using an eigenvalue criterion of 1.0, three factors were extracted. The three common factors accounted for 57.22% of the total task variance. From the loadings on the varimax-rotated factors, presented in Table 3, the following names were given to the factors: (1) FI-Memory Accessibility, where the highest loadings were for *Relationship Eduction* and *Short-term Retention,* (2) FII-Response Implementation, where the highest loadings were for *Response Decision* and *Response Recording,* and (3) FIII-Decision, with the highest loadings for *Relationship Evaluation* and *Relationship Study. Analogy Completion After Study* loaded moderately only on Memory Accessibility, while *Relationship Choice* and *Relationship Evaluation* split loadings on Memory Accessibility and Decision. *Choice Reaction Time* did not load significantly on any factor.

Table 3 presents data on the accuracy of task

Table 3
Means, Standard Deviations and
Factor Loadings for Task Responses

Variables	$\bar{X}$	SD	Factor I	Factor II
Response Decision	.89	.24	.50	-.56
Short-term Retention	.89	.11	.32	-.02
Analogy Completion After Study	.36	.18	.73	.02
Relationship Eduction	.95	.08	.21	.41
Relationship Evaluation	.58	.17	.72	.10
Relationship Choice	.56	.20	.66	.17
Proportion of Common Variance			.78	.22

responses. The *Choice Reaction* data is not given, as no errors were made on the task. It can be seen that, except for *Analogy Completion After Study,* the mean proportion of correct responses on the tasks was substantially higher than that for solving the 27 analogies presented in the standard multiple-choice format. *Response Decision, Short-term Retention* and *Relationship Eduction* were easy tasks. *Relationship Evaluation* and *Relationship Choice* were somewhat more difficult, but not as difficult as the standard analogy task. A principal factor analysis was performed on the task responses, with squared multiple correlations as initial communality estimates. Using an eigenvalue criterion of 1.0, two factors were extracted, which accounted for 40% of the total task variance. After varimax rotation, only one factor accounted for a substantial proportion of the common variance. An inspection of the loadings suggested that the first factor concerned evaluation, while the second factor seemed uninterpretable.

Factor scores were obtained for both the response-time and response-accuracy factors. Since common factor analysis, rather than principal component analysis, was applied to the task correlations, factor scores must be estimated. The factor score coefficient matrix was obtained from the product of the factor structure matrix with the inverse of the correlation matrix. The individual factor scores were obtained from the product of the factor score coefficient matrix and the matrix of z scores for subjects on the task components. Factor scores for subjects with missing data were obtained by substituting the missing task mean.

Table 4 presents the regression of the analogy item performance (from the *Analogy Test*) separately on the task component response-time and response-accuracy factors. The multiple correlation of number of analogies solved (*Analogy Test* score) on the task response time factors was moderate ($R = .50$), with Factor I (Memory Accessibility) and Factor III (Decision) showing significant negative and positive beta weights, respectively. Additionally, the multiple correlation of analogy solving latency with the task latency factors was moderately high for both correct and incorrect items. All the factors had significant weights in predicting analogy solving latency for correct items, with Factors I and III receiving approximately equal positive betas, and Factor II (Response Implementation) receiving a small negative beta weight. For the incorrect items, Factor III received a large positive beta while Factor I received a small positive beta.

The multiple correlation of the number of analogies solved with the task accuracy factors was also moderately high ($R = .66$), with only Factor I (Evaluation) showing a significant cor-

Table 4
Regression of Analogies on Task Factors

Variable	Response Time Factors				Response Accuracy Factors		
	Betas			Multiple	Betas		Multiple
	FI	FII	FIII	R	FI	FII	R
Analogy Test	-.41**	-.15	.29**	.50**	.66**	.02	.66
Analogy Latency (Correct Items)	.37**	-.21**	.43**	.64**	-.20	-.19	.27
Analogy Latency (Incorrect Items)	.23*	-.05	.59**	.67**	.06	-.01	.06

*
 Statistically significant at p≤.05
**
 Statistically significant at p≤.01

relation. The multiple correlations of the analogy solving latencies with the task accuracy factors were not significant ($p > .05$).

To examine the influence of the task performance factors on the relationship between achievement and analogy solving, *Analogy Test* scores were decomposed into predicted and residual scores from each of the regressions on the component tasks. From the response-time factors, the predicted *Analogy Test* score correlated .24 with GPA ($p < .05$), while the residual score correlated .38 with GPA ($p < .01$). Similarly, when the response-accuracy factors were used as predictors, the predicted *Analogy Test* score correlated .37 with GPA ($p < .01$), while the residual score correlated .28 ($p < .05$). Thus, although some of the correlation between GPA and analogy solving may be accounted for by information-processing factors, a substantial proportion of variance in GPA which is explained by analogy item-solving is not predictable by the factors.

Discussion

The results generally suggest the feasibility of explaining individual differences in performance on a prototypic intelligence test item, verbal analogies, from the success and efficiency of processing information on some hypothesized component tasks in item solving. The two separate measures recorded on the set of component tasks—response accuracy and response time—will be discussed separately since these variables showed different patterns of relationships in the current study, as well as have generally different theoretical implications.

Response Time

The response time data suggest that at least three types of information-processing events are needed to describe individual differences in the component task durations. A factor analysis of the component task correlations identified three common factors in the response-time data. An inspection of the task loadings suggested that the following names described the orthogonal factors: (1) Memory Accessibility, which was loaded on by both long- and short-term memory tasks; (2) Response Implementation, a factor which is loaded on by a low-level matching task and resonse recording and (3) Decision, a factor involving evaluation and choice tasks.

Interestingly, the three orthogonal factors obtained under analytic rotation (varimax, in this case) corresponded fairly well to the a priori classification of tasks into three areas of information-processing events—memory processes, response processes and control processes, respectively. However, results of the factor analysis should be regarded as suggestive rather than as a hypothesis test for these areas of information processing, since the number of component events was small, and the factoring was done by the traditional exploratory method rather than the more advanced confirmatory methods. More importantly, at this stage of research on analogy items, the emergence of three factors for response latencies indicates that even if task outcome is controlled across individuals (i.e., the correct response is obtained), substantial individual differences in the patterns of mental processing remain.

An important issue for this study is the extent to which the component task factors account for individual differences in the total duration of processing in solving analogies. The results indicated that the response time factors accounted for analogy-solving latencies moderately well ($R = .64$), with Memory Accessibility and Decision showing significant positive beta weights. Additionally, Response Implementation showed a smaller, but significant, negative beta weight. Although the explanation for the reversed direction of Response Implementation is not immediately apparent, the magnitude of the correlations show that the task component factors have at least some potential to account for individual differences in information-processing duration on analogies.

More significant in the present study is the extent to which individual differences in general ability, as measured from solving analogies, re-

sults from differences in mental processing. *Analogy Test* scores were moderately correlated with the task latency factors ($R = .50$); Memory Accessibility and Decision had significant negative and positive correlations, respectively. These results suggest that the most successful analogy solver, and consequently the individual of highest ability, encodes and retrieves memory information about the items rapidly, but spends relatively more time evaluating the relational quality of the information retrieval. The results on memory accessibility support the findings of Hunt et al. (1973, 1975) on the information-processing implications of high IQ scores. However, the results on complex decision processes have no direct parallel in other studies.

In contrast, total item solving latency does not correlate significantly with success in solving items ($r = -.06$). This finding is consistent with earlier psychometric studies on item accuracy versus item rate (Baxter, 1941; Davidson & Carroll, 1945), but, on the surface, appears inconsistent with the predictability of item solving from the mental processing latency factors. The explanation is that success in solving items (hence ability) is not associated with *total speed* of information processing, but apparently depends on the *relative duration* of the various component events.

Response Accuracy

The response accuracy data show a different pattern of findings for the component tasks. The relative difficulties of the tasks are especially important in these data, as they can be interpreted as indicating the probability of successfully completing a component. The data showed that the response task and the memory tasks were successfully completed quite regularly, while the tasks involving higher-order evaluation processes were fairly difficult. Thus, on the basis of difficulty alone, the response and memory tasks show little promise in accounting for individual differences in item solving.

The factor analysis and the regression data for task accuracy showed a similar lack of potential for the response and memory tasks. Two common factors emerged, but only the factor which was loaded on by the evaluation and judgment tasks accounted for much common variance or correlated with analogy solving. The response task split loadings between the two factors, and while one memory task *(Relationship Eduction)* had a marginal loading on the small factor, the other memory task *(Short-term Retention)* loaded on no common factor.

On the surface, the findings from the response-time data and response-accuracy data seem somewhat inconsistent. The response-time regression data revealed that the memory factor was the most important predictor of analogy-solving ability. Since the primary component task loading on the memory factor was *Relationship Eduction,* i.e., retrieving word relationships from memory, this finding is consistent with previous literature which indicated the importance of relationships in solving analogies (Whitely, in press, 1976; Whitely & Dawis, 1974). Yet, the response-accuracy data show *Relationship Eduction* as having little promise of explaining individual differences in item solving.

The resolution to these inconsistencies most likely involves the nature of the word relationships which are retrieved from memory. In the present study, response accuracy for relationship retrieval was scored when the subject produced a meaningful relationship between word pairs, outside the context of any specific analogy. However, in an analogy item, several relationships may be retrievable for a given word pair, but only one of these relationships may be appropriate to the item. Although the task accuracy data show that simply producing any relationship is not important in explaining individual differences in solving analogies, the response time data show that speed in word relationship retrieval is correlated with analogy-solving ability. A possible explanation for the inconsistent findings is that the higher speed of retrieving word relationships leads to a greater likelihood of educing the appropriate relationship in analogies. Further research on this hypothesis would appear profitable, with special

emphasis on the underlying mechanisms for relationship eduction.

A final conclusion that may be drawn from the results is that a more complex account of solving analogies is needed than is possible from the components measured in the current study. An adequate explanation of individual differences on analogy tests should result in high predictability of both analogy-solving success and response time. The multiple correlations between performance on psychometric analogies and task components were only moderate, and substantial proportion of variance in school achievement (GPA), which was correlated with analogy-solving success, was not predicted by the component factors. Although anticipating the nature of more complete explanations is admittedly speculative at this point, it would seem that individual differences in item-solving strategies and the content of memory stores should receive more systematic attention. The latter probably determines the nature of the word relationships which are retrieved, while the former may produce variability in both response accuracy and response time, which would not be accounted for by individual differences in component events, considered separately.

References

Baxter, B. An experimental analysis of the contributions of speed and level to an intelligence test. *Journal of Educational Psychology*, 1941, *32*, 285–296.

Allison, R. *Learning parameters and human abilities.* (Technical Report). Princeton, NJ: Educational Testing Service, 1960. (ONR Contract 694 (001), NSF Grant G-642)

Carroll, J. B. *Psychometric tests as cognitive tasks: A new structure of intellect.* Princeton, NJ: Educational Testing Service, May 1, 1974.

Davidson, W. M., & Carroll, J. B. Speed and level components in time-limit scores: A factor analysis. *Educational and Psychological Measurement*, 1945, *5*, 411–427.

Dunham, J. L., Guilford, J. P., & Hoepfner, R. Multivariate approaches to discovering the intellectual components of concept learning. *Psychological Review*, 1968, *75*, 206–221.

Hunt, E., Frost, N., & Lunneborg, C. Individual differences in cognition: A new approach to intelligence. *The Psychology of Learning and Motivation.* New York: Academic Press, 1973, *7*, 87–122.

Hunt, E., Lunneborg, C., & Lewis, J. What does it mean to be high verbal? *Cognitive Psychology,* 1975, *7*, 194–227.

Lunneborg, C. E. *Individual differences in memory and information-processing.* Seattle: University of Washington, Educational Assessment Center, 1974.

Lunneborg, C. E. Choice reaction time: What role in ability measurement? *Applied Psychological Measurement,* 1977, *1*, 309–330.

Messick, S. Beyond structure: In search of functional models of psychological process. *Psychometrika,* 1972, *37*, 357–375.

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

Tinsley, H. E., & Dawis, R. V. An investigation of the Rasch simple logistic model: Sample free item and test calibration. *Educational and Psychological Measurement,* 1975, *35*, 325–340.

Whitely, S. E. Relationships in analogies: A semantic component of a psychometric task. *Educational and Psychological Measurement,* in press.

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

Whitely, S. E., & Dawis, R. V. Effects of cognitive intervention on latent ability measured from analogy items. *Journal of Educational Psychology,* 1974, *66*, 710–717.

Author's Address

Susan E. Whitely, Psychology Department, University of Kansas, Lawrence, KS 66044