DOCUMENT RESUME

ED 194 639 TM 800 772

AUTHOR Hunt, Earl B.; And Others
TITLE Individual Differences in Long-Term Memory Access.
PUB DATE [79]
NOTE 56p.; Broken type may be marginally legible.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Classification; *Cognitive Processes; Higher Education; *Individual Differences; *Memory; *Reaction Time; Reading Comprehension; *Semantics; Verbal Ability
IDENTIFIERS Nelson Denny Reading Tests

ABSTRACT

In a study of individual differences in long-term memory access, university undergraduates verified (1) whether an item was a member of a category; (2) whether two items belonged in the same category; and (3) whether two words had the same name. Reaction times from these tasks were correlated with verbal ability, as measured by performance on a standardized test of vocabulary and reading comprehension. A relationship was found between verbal ability and rection time in the verification tasks. These results are contrasted with those of Hosaboam and Pellesrino (1978), who failed to find such a relationship. The results indicated that the varous reaction time measures form a single factor that bears a moderate relationship to reading and vocabulary measures. (Author/GK)

INDIVIDUAL DIFFERENCES IN LONG-TERM MEMORY ACCESS

Earl B. Hunt, Janet Davidson, and Marcy Lansman
University of Washington

Short title: Individual Differences in Memory Access

Send Proofs to: Dr. Earl B. Hunt
Department of Psychology, NI-25
University of Washington
Seattle, Washington 98195

Abstract

In a study of individual differences in long-term memory access, university undergraduates verified a) whether an item was a member of a category, b) whether two items belonged in the same category, and c) whether two words had the same name. Reaction times from these tasks were correlated with verbal ability, as measured by performance on a standardized test of vocabulary and reading comprehension. A relationship was found between verbal ability and reaction time in the verification tasks. These results are contrasted with those of Hogaboam and Pellegrino (1978), who failed to find such a relationship. The results indicated that the various reaction time measures form a single factor that bears a moderate relationship to reading and vocabulary measures.

Written language depends on the arbitrary connection of symbols with their referents. There is no particular reason why the symbols "A" and "a" should refer to the first letter of the alphabet. Similarily, there is no reason why "cat" should be the linguistic symbol for a small, domesticated feline. We simply have to memorize the associations that our language uses. In the last ten years there have been several studies indicating that the speed of retrieving such associations is a stable dimension of individual differences, and that this dimension bears a moderate, but reliable, statistical relationship to scores on conventional written aptitude tests. The correlation is -.3 (long retrieval times are associated with low aptitude) in homogeneous groups of subjects, such as college students. The absolute differences in retrieval times may be considerable if one compares groups of widely varying abilities, such as normal young adults and educable mental retardates (Hunt, 1978, 1980).

Much of the evidence for the assertion that name retrieval is related to verbal aptitude is based upon studies using some variant of the stimulus matching paradigm developed by Posner and Mitchell (1967). In a stimulus matching paradigm, the subject is presented with two visual symbols, and asked to indicate whether or not they name the same item. Examples are the letter pair A-A, which is a physically identical (PI) pair, and A-a, which is a name identical (NI) pair. Responses are faster to PI than to NI pairs, and the difference between the two types of pairs is smaller in groups of higher verbal ability. Indeed, both name matches and physical matches tend to be faster in higher ability groups, although time for a name match is more highly correlated with verbal ability measures than is time for a physical match (Hunt, Lunneborg, & Lewis, 1975).

The fact that performance in the stimulus matching paradigm is tied to verbal aptitude is not, in itself, of great interest. Interest arises from the interpretation of the paradigm. If reaction time in the stimulus matching paradigm is a measure of speed of retrieval of overlearned information, then the association of retrieval speed with verbal aptitude gives us information about the processing which underlies a complex talent. But is this interpretation of the task appropriate? As is virtually always the case in experimental psychology, several models have been proposed to explain performance in this task (Posner, 1978). Thus in order to buttress the theoretical conclusion, it would be desirable to show that other measures of the speed of access to overlearned semantic information are correlated with verbal ability.

Hogaboam and Pellegrino (1978) attempted to do just this. They used a semantic categorization task in which observers were first told a category name, and then presented with a sequence of nouns and line drawings. The task was to indicate whether each item in the sequence was an exemplar or non-exemplar of the category. Hogaboam and Pellegrino found (as had other investigators) that reaction time for categorization was a function of stimulus type, frequency of occurrence of the stimulus as an exemplar of its category, and type of decision (positive or negative). However, they did not find any relationship between category identification time and verbal ability, as measured by the Scholastic Aptitude Test (SAT). Hogaboam and Pellegrino concluded that previous findings based on the stimulus matching paradigm might not be due to individual differences in the process of retrieving a name from memory. They argued instead that the previous findings might have been due to more intelligent subjects being more adept at responding to the unusual task demands of the stimulus matching paradigm.

Hogaboam and Pellegrino raised an important issue. If the results obtained from the stimulus matching paradigm are due to task demands, rather than to the retrieval of information from memory, the previous results should be reinterpreted. (Indeed, the phenomemon of individual differences in adjustment to task demands would itself be worth explanation.) Before drawing this conclusion, however, we wished to replicate the empirical findings reported by Hogaboam and Pellegrino. In particular, certain aspects of their statistical treatment and their interpretation of the semantic categorization paradigm raised some questions.

To summarize Hogaboam and Pellegrino's argument, they observed that a) measures derived from the semantic categorization task were sufficiently reliable to produce the nomothetic effects (e.g. effects of taxonomic category and frequency) observed by others, but b) these measures were not "significantly" correlated with measures of verbal ability.

They concluded that because the semantic categorization task certainly requires subjects to access semantic information in long-term memory, the failure to find significant correlations is evidence against a relationship between speed of access to semantic information and verbal ability.

Three issues cloud interpretation of Hogaboam and Pellegrino's data: reliability of their measures, power of their design, and theoretical interpretation of their semantic categorization paradigm. The fact that Hogaboam and Pellegrino found several experimental effects to be statistically reliable is irrelevant to the question of whether their reaction time measures provided reliable measures of individual performance. The statistical tests required to establish nomothetic effects are different from the statistical tests required to establish individual differences effects, and it is quite possible for a set of data to be reliable with respect to one type of test and unreliable with respect to the other. Hogaboam and Pellegrino did not collect data on the reliability of their measures as indicants of individual performance.

A second cause for concern is that Hogaboam and Pellegrino used only 34 subjects in their experiment. Thus their study might be faulted for a lack of statistical power, since the statistical power of a test of the null hypothesis that $r = 0$, given a true correlation of -.3, is only .54. An alternative to testing the hypothesis that $r = 0$ is to use a likelihood ratio to compare the probability that a given correlation coefficient was obtained by sampling from a population in which $r = 0$ to the probability that it was obtained by sampling from a population in which r = -3. In their published report, Hogaboam and Pellegrino did not present the correlations they obtained, but Pellegrino has since provided them for us. They range from 0 to -.25. In most cases the liklihood ratios favor the hypothesis that the sample was drawn from a population in which r=0.

Schwartz (1980) raised a non-statistical question about Hogaboam and Pellegrino's experiment. The procedure required the subject to memorize a category name (e.g. BIRD) and then determine whether 24 stimuli were exemplars or non-exemplars of that category. Thus the subject was asked to hold the category name in short-term memory while responding to the items to be categorized. This differs from the stimulus matching procedure, in which short-term memory requirements are much reduced. It has been observed that short-term memory performance does not correlate with verbal ability scores, except when the contrast is between widely separated groups such as university students and mental retardates (Hunt, 1978; see also Matarazzo, 1972 for a discussion of some psychometric evidence.) Schwartz argued that short-term memory requirements in the semantic categorization task may have affected the observed pattern of individual differences. Another complicating fact is that some of Hogaboam and Pellegrino's measures involved the categorization of pictures. Whether or not the naming of a picture is the same as the naming of a word is at least arguable. Indeed, some would claim that different brain regions are involved (Walsh, 1978).

Because Hogaboam and Pellegrino did not observe correlations between verbal ability and semantic categorization, they concluded that the correlation between verbal ability and stimulus matching was due to special aspects of the stimulus matching task that do not involve access to name codes. An alternative explanation is that the failure to observe correlations between verbal ability and the semantic categorization task might be due to the introduction of information processing requirements into the semantic categorization task that do not appear in stimulus matching and that are not associated with verbal ability.

In order to resolve these issues, Hogaboam and Pellegrino's study was replicated and extended. The number of subjects was increased to address the issue of statistical power. In addition, the number of trials contributing to each measure was increased, and data was collected on the reliability of the measures. Finally, the experiment was designed to examine the relationship between short-term memory and verbal ability.

Method

The experiment consisted of three phases. In the first phase, Hogaboam and Pellegrino's experimental procedure was replicated as exactly as possible, based on the information in their published report. In the second phase, subjects completed additional trials in the Hogaboam and Pellegrino semantic categorization task in order to increase the reliability of the reaction time measures. Finally, a number of other experimental tasks were administered in an effort to examine more carefully the differences between the semantic categorization and the stimulus matching paradigms. In these additional tasks, the short-term memory demands and the nature of the yes-no decision were both varied.

Hogaboam and Pellegrino used the verbal score from the SAT as their criterion measure of verbal ability. Some of our subjects had verbal scores on the Washington Pre-College (WPC) Test, which is similar to the SAT. However, since college entrance examination scores were not available for all of our subjects, our criterion measure was the Nelson-Denny Reading Test (1960), which provides measures of reading rate, vocabulary, and reading comprehension. High correlations have been found between the Nelson-Denny Test and verbal scores on the WPC (Palmer, MacLeod, Hunt, and Davidson, Note 1).

Subjects

The subjects were 75 University of Washington undergraduates. They were paid $15 for their participation. Subjects were run in groups of 2 to 6. Due to attrition and computer failure, data was not obtained on all measures from all subjects.

Apparatus

A NOVA 3 computer controlled stimulus presentation and response collection for the reaction-time tasks. During Session 1, which involved presentation of both words and pictures, subjects were seated at individual desks that held response keyboards. Slides were projected by a Kodak carousel 850H slide projector onto a 8 x 8 foot white screen viewed by all subjects in the group. Pioneer SE20A 8I headphones were used for auditory feedback. During Sessions 2 and 3, subjects were seated in individual sound-attenuating booths. Each booth contained a response keyboard and an independently-controlled Tektronix 604 cathode ray tube oscilloscope on which the stimuli were presented.

Procedure

Subjects participated in three one and one-half hour sessions. All subjects performed the same tasks in the same order. Table 1 shows the order of task presentation.

Insert Table 1 about here

For all tasks, subjects were told to work as quickly as possible without making errors. Feedback was provided for the reaction-time tasks. During Session 1, subjects heard a tone whenever they made an error. During Sessions 2 and 3, subjects received "OK" or "No" messages on their screens, after each trial, informing them of the accuracy of their response. In addition, during Sessions 2 and 3, after every set of eight trials subjects saw their reaction-times in milliseconds averaged over those eight trials, and the total number of errors they had made so far. If they wished, subjects could relax for a moment in their booths at this time. When they were ready to begin the next set of eight trials, they pressed any key on their response boards and the next trial began 250 msec later. Thereafter, trials were spaced 500 msec after trial feedback. Every trial was preceded by a warning dot which appeared for 250 msec in the center of the screen. There was a 250 msec interval between the offset of the warning dot and the onset of the stimulus.

SESSION 1

Semantic Categorization. This task, which was a replication of Hogaboam and Pellegrino's procedure, was performed at the beginning and end of Session 1. Subjects were asked to determine whether an item belonged to a specified semantic category. For example, they might be asked to respond as to whether a car is a vehicle. As in the Hogaboam and Pellegrino study, the furniture category was used for practice trials, and the experimental categories were: carpenter's tools, vehicles, body parts, four-legged animals, weapons, articles of clothing, kitchen utensils, musical instruments, insects, and fruits. Three high and three low taxonomic frequency items from each of the experimental categories were selected from the Battig and Montague (1969) norms. The average taxonomic frequency was 308 for the high frequency items and 32 for the low frequency items. (Hogaboam and Pellegrino's average frequencies were 350 and 34, respectively.) The Kucera and Francis (1967) average printed frequency was 64.1 for the high frequency items and 3.3 for the low frequency items. (Hogaboam & Pellegrino's average printed frequencies were 84 and 4.5 for high and low items, respectively.) Each item was depicted as an unambiguous line drawing and as an uppercase printed word. Slides were made of the drawings and words. Slides were also made of each category name printed in uppercase letters and underlined.

A block of trials consisted of a category name followed by 24 items. Twelve of the items were exemplars and 12 were not. The positive items (exemplars) were three high and three low taxonomic frequency members of the category, each presented once as a word and once as a picture. Negative items for the practice block were taken from categories not used in this experiment. Negative items for the first five categories were taken from the positive items of the last five categories. Negative items for the last five experimental categories were taken from the positive items of the first five categories. The 12 negative items in each block consisted of three items of high taxonomic frequency in their own categories, and three items of low taxonomic frequency. Items were represented as both pictures and words. Half of the subjects saw one set of five categories first, and the other half of the subjects saw the other set of five categories first. The order of categories within each set of five categories was varied for each group of subjects according to a Latin square design.

Each category name was displayed for 3000 msec. Each item was displayed for 1000 msec. Subjects were instructed to press the right key if an item was a member of the category and the left key if it was not. Subjects heard a tone through headphones whenever they made an error.

Nelson-Denny Vocabulary Test. Form B of the vocabulary portion of the Nelson-Denny (1960) test was administered. Subjects were given ten minutes to answer 100 items. Subjects were instructed to select the best definition of a word from among five choices. Responses were recorded on the Nelson-Denny self-marking answer sheet.

Nelson-Denny Reading Test. Form B of the comprehension and reading rate portion of the Nelson-Denny (1960) test was administered. Subjects were given 20 minutes to read eight selections and answer four multiple-choice questions at the end of each selection. While answering questions, subjects were allowed to refer back to the relevant passage. If they had difficulty with a question, they were told to make a reasonable effort and then go on to the next question. Subjects used the Nelson-Denny self-marking answer sheet to record their responses.

The first minute of the test was used to determine reading rate. Subjects began reading the passage and, at the end of one minute, they were told to stop on the line they were reading and record that line number on their answer forms. When this was completed, subjects resumed their reading.

Sentence-Picture Comprehension Test. This paper and pencil test is divided into five sections (Lansman, 1977). Because of time restraints, only the first three sections were administered. Section 1 was considered practice. Each section contains 64 items. For each item, subjects were asked to determine whether a sentence accurately described a picture (Clark & Chase, 1972). The stimuli consisted of 16 descriptive sentences that varied in linguistic complexity (e.g. plus above star; star isn't above plus), and two pictures ($\begin{smallmatrix}+ & * \\ * & +\end{smallmatrix}$). If the sentence correctly described the picture, subjects were to mark the T for true; otherwise, they were to mark the F for false.

Subjects were given 2.5 minutes for each of the three sections.

The sentence-picture comprehension task provided a logical verbal task that, presumably, depends less upon access to the meaning of words than the other tasks. Obviously, a person doing a sentence-picture comprehension task must know what the words mean. However, only six words are used in this task. Thus it is more dependent upon the manipulation of verbal items in working memory (Baddeley, 1976) than upon the retrieval of information about the names of the stimuli. In a previously conducted analysis of the information processing correlates of reading, we found that sentence verification and stimulus matching were predictors of different components of reading skill: sentence-picture comprehension was more related to reading comprehension and stimulus matching was more related to reading speed (Palmer et al., Note 1). Thus, we might expect the semantic tasks and the sentence-picture comprehension task to make independent contributions to the prediction of verbal ability scores.

SESSION 2

Two-choice reaction time. A simple choice reaction time task was given in order to familiarize subjects with the apparatus and to obtain a basic measure of key-pressing time. On each trial of this task, subjects saw three stars on the screen to either left or right of center. They pressed the RIGHT key if stars were on the right, and the LEFT key if stars were on the left. There were eight practice trials and 48 experimental trials.

Stimulus (word)-matching. This task required subjects to determine whether two words had the same name (Palmer et al., Note 1). The stimuli were pairs of common four-letter words. The "different" trials consisted of word pairs differing by one letter (e.g. DATE GATE), with the changed letter balanced over position. The words were presented in upper and lower case. There were 24 physically identical pairs (DATE DATE), 24 name identical pairs (DATE date), and 48 different pairs (date gate). Each pair was shown twice. Subjects were to respond SAME by pressing the right key if the words had the same name, and to respond DIFFERENT by pressing the left key if the words had different names.

In the simultaneous condition of this task, the words appeared side by side on the screen with one character space between words. The word pairs remained on the screen until the subject responded. There were 16 practice trials and 192 experimental trials of this condition.

In the sequential condition, devised to add a short-term memory component, one word appeared on the screen for 500 milliseconds. Then, 1500 milliseconds later, the second word appeared and remained on the screen until the subject responded. There were 16 practice trials and 192 experimental trials.

Semantic Verification. This task, like the semantic categorization task, required subjects to determine whether an item was a member of a category. A category name and an item were presented on each trial. All stimuli were presented as uppercase words. Subjects were to press the right key if the item was a member of the category; otherwise they were to press the left key.

The category names and items used in this task were a subset of those used in the semantic categorization task. The eight experimental categories were: furniture, body part, fruit, weapon, clothing, insect, vehicle, and animal. Three instances of high taxonomic frequency and three instances of low taxonomic frequency were selected from each category. Each instance appeared once as a positive item and once as a negative item.

In the simultaneous condition, the category name and the item appeared on the screen at the same time. The category name was directly above the item and both pieces of information remained on the screen until the subject responded. There were 16 practice trials and 96 experimental trials.

In the sequential presentation condition, the category name appeared on the screen for 500 milliseconds, and 1500 milliseconds later the item appeared. The item remained on the screen until the subject responded. There were 16 practice trials and 96 experimental trials.

Semantic Matching. The semantic categorization and the semantic verification tasks involve making decisions about superordinate-subordinate (category-instance) relationships, whereas the stimulus matching task involves making a subordinate-subordinate or "instance-instance" decision. The semantic matching task required the subject to make a semantic decision about two instances. On each trial of this task, subjects had to decide whether two items were members of the same category (Goldberg, Schwartz, & Stewart, 1977). Subjects were to press the right key if the two items belonged in the same category (APPLE PEACH); otherwise they were to press the left key (APPLE PLANE). All stimuli were presented in uppercase letters. Four instances of high taxonomic frequency and four instances of low taxonomic frequency were selected from the following categories: vehicle, fruit, body part, furniture, insect, and clothing. Each instance appeared twice as a positive item and twice as a negative item. Instances of high taxonomic frequency were paired together and instances of low taxonomic frequency were paired together: there were no instances of high taxonomic frequency paired with instances of low taxonomic frequency.

In the simultaneous presentation condition, two items appeared on the screen at the same time, with one item centered directly above the other. Both items remained on the screen until the subject responded. There were 16 practice trials and 96 experimental trials.

In the sequential presentation condition, one item appeared on the screen for 500 msec and 1500 msec later the second item appeared. The second item remained on the screen until the subject responded. There were 16 practice trials and 96 experimental trials.

SESSION 3

Subjects performed the same tasks during Session 3 that they performed during Session 2. The two-choice reaction time task was presented first and the others were presented in reverse order from Session 2.

Results

The results section is divided into three parts. The first part is a general description of the data. The second presents the group results for each task. The third part is a discussion of the individual differences results.

General description

The mean reaction time for each trial type was calculated for each subject. Data were analyzed only for trials on which the response was correct and on which reaction times were within three standard deviations of the mean for that subject in that condition.

Split-half reliabilities (odd versus even trials) were calculated for correct reaction times and corrected using the Spearman-Brown formula. All mean reaction time measures were highly reliable , with reliabilities ranging from .94 to .99.

Each reaction time task was presented twice. Because the same experimental effects were observed on each occasion, data from the two presentations were combined for each task, with the exception of the semantic categorization task, in order to increase the number of observations.

Group Results

Semantic Categorization. The semantic categorization task was presented twice, once in an exact replication of Hogaboam and Pellegrino's procedure, and then again to increase the reliability of the data. For each presentation of this task, mean reaction times were calculated for each subject for each combination of stimulus type (words, pictures), taxonomic frequency (high, low), and decision type (yes, no). Table 2 shows the mean reaction times for these measures.

Insert Table 2 about here

The results from both presentations of this task replicated the main results found by Hogaboam and Pellegrino (1978). Positive decisions were faster than negative decisions, category membership of pictures was verified faster than membership of words, and items of high taxonomic frequency were verified faster than items of low taxonomic frequency ($p < .001$ for these comparisons).

Semantic Verification (FRUIT APPLE) and Semantic Matching (APPLE PEACH). The results from each of these tasks were similar to those found in the semantic categorization task. Positive decisions were faster than negative decisions, and items of high taxonomic frequency were verified faster than items of low taxonomic frequency ($p < .001$ for these comparisons). The mean reaction times from the semantic verification task and the semantic matching task are shown in Table 3.

Insert Table 3 about here

Stimulus Matching (DATE data). Reaction times from this task are shown in Table 4. In the simulataneous condition, mean PI reaction time was 121 msec faster than mean NI reaction time. Differences of approximately this size have been found in previous studies. In the sequential condition, the difference between NI and PI reaction times was much smaller (19 msec). The finding of a smaller difference between NI and PI reaction times when stimuli are presented sequentially is also a very common result. It has been attributed to the fact that the physical match, as well as the name match, must be made on the basis of the names of the stimuli, since the visual trace of the initial item has faded by the time the second item appears (Posner, 1978).

Insert Table 4 about here

Two-Choice Reaction Time. The mean reaction time for right hand responses on correct trials was 276 msec. The mean reaction time for left hand responses on correct trials was 287 msec.

Paper and Pencil Tests. Table 5 shows the mean scores for number correct, standard deviations, and range of scores for the Nelson-Denny Vocabulary, Nelson-Denny Reading Comprehension, and sentence-picture comprehension tests.

Insert Table 5 about here

Individual Differences.

Hogaboam and Pellegrino reported eight processing speed measures: the time to categorize items representing the eight possible combinations of high and low frequency, pictures and words, and positive and negative decisions. In our data, the reliabilities of these measures were high. Split-half reliabilities, based on odd versus even trials, ranged from .92 to .97. Table 6 shows the correlations between these eight measures and the four measures taken on the Nelson-Denny test: vocabulary, paragraph comprehension, reading rate, and a composite score (Nelson and Denny's suggested composite = 2 * comprehension + vocabulary). The left half of Table 6 shows the correlations based on the same number of trials used by Hogaboam and Pellegrino. The right half shows correlations based on RTs from both presentations of the semantic categorization task. Thus the left half of the table can be considered a replication of the Hogaboam and Pellegrino procedure, and the right half a replication using twice the number of trials.

Insert Table 6 about here

Consider first the left half of the table. The eight correlations of RT measures with the Nelson-Denny composite range from -.14 to -.31. Six of the eight are statistically significantly different from 0, at the .05 level. The correlations based on additional data are shown in the right half of the table. Seven of the eight are statistically significant. Clearly these results are more consistent with the hypothesis that the population correlation is -.3 than with the hypothesis that it is zero. This statement can be quantified by calculating the following likelihood ratio: (probability of r given that population r = -3) divided by (probability of r given that population r = 0). For r = -.25 and n=67, typical of data in Table 5, the ratio is greater than 6:1 in favor of the hypothesis r = -.3.

Hogaboam and Pellegrino also computed eight difference scores which represented word-picture differences for high- and low-frequency, positive and negative items, and low-high frequency differences for positive and negative pictures and words. The relevant results from our study are shown in Table 7, which is similar in format to Table 6. The reliabilities of the difference scores were low, and one is even negative! With such low reliabilities, it is hardly surprising to find that there were few significant correlations with the Nelson-Denny measures.

Insert Table 7 about here

Reliability aside, the theoretical reasoning behind the use of some of the difference scores as correlates of verbal ability is not clear. Consider the contrasts based upon picture versus word reaction times. If the same processes are involved in semantic categorization of pictures and words, then the interpretation of an individual difference measure based upon differences is not meaningful. If, as could be argued, the two RTs are based on different processes, then a difference between the two is only of interest if the processes required by one are a subset of those required by the other. This is the case for at least some models of the stimulus matching task, where the processes involved in producing RTs to PI pairs are contained within the processes involved in reacting to NI pairs (Posner, 1978). However, we know of no such model for picture versus word categorization.

There is also a problem in interpreting negative-positive difference scores. Presumably negative responses are made if some measure of the semantic distance between the category name's representation and the target stimulus is exceeded. Unless this criterion is different for high as opposed to low frequency items, or for pictures versus words, one would expect the negative semantic comparisons to have similar latencies. This would produce low reliabilities for difference scores, since each component of the difference would be based largely upon a measure of the same process. The lowest reliability estimates for the measures (and the only reliability measures not reliably greater than zero) were obtained for difference scores involving negative responses.

This leaves us with two contrasts: one based on the difference between positive RTs to high and low frequency words, and one based on the positive response RTs to high and low frequency pictures. It is reasonable to speculate that "low" verbal people might be as fast as "high" verbal people in accessing information about high frequency words, but would be slower to access low frequency words. Our results show this to be the case. The conclusion is slightly more interesting because the same finding is true for pictures. The relevant data are shown in Table 7. Fortunately, the reliability coefficients for these difference scores are above chance, although they are still low.

In summary, the individual differences data presented here would not lead one to the negative conclusion drawn by Hogaboam and Pellegrino. A relationship was found between category identification time and verbal ability. There is, however, one major difference between our study and theirs. We were forced to use the Nelson-Denny test as our major criterion measure. Only a small number of subjects (26) who completed our replication of the Hogaboam and Pellegrino procedure had also taken the Washington Pre-College Test, a test similar to the

SAT used by Hogaboam and Pellegrino (Footnote 1).

Short-term memory: A test of Schwartz's Hypothesis. S. Schwartz (1980) observed that the semantic categorization task contained a short-term memory component; the subject must remember the category name while responding to the stimuli to be categorized. Schwartz noted that short-term memory seems to be only loosely related to psychometrically defined verbal ability. He suggested that the short-term memory component of the categorization task may have introduced additional variance that masked the relation between access to information in long-term memory and verbal ability. Schwartz's hypothesis could be tested by manipulating the short-term memory demands of the task and observing whether or not the correlations with psychometric tests are affected. Our experiment was designed to do this. In the sequential version of each task, the subject was required to hold the first item in memory while responding to the second. This introduced a minimal short-term memory requirement into the task. In the simultaneous versions, the short-term memory component was eliminated.

This short-term memory manipulation did not change the pattern of correlations. Table 8 shows the correlations between the verbal measures and RTs to positive responses for word items in the semantic categorization task and for positive responses in the simultaneous and sequential versions of the other semantic tasks. Only positive responses were analyzed because, as mentioned, the models for negative responses are unclear. The table also shows the correlation between Nelson-Denny scores and the NI-PI reaction time difference score computed for simultaneous and sequential versions of the stimulus matching task (Date date). The correlations of all tasks with the Nelson-Denny composite are all slightly below .3 in absolute magnitude, and do not vary in any consistent way with the amount of short-term memory involved.

The correlations between the RT measures and the Nelson-Denny reading rate present a different picture. Only those tasks involving simultaneous presentation of the items had significant correlations with reading rate. This is not surprising because subjects were timed for reading more material in the simultaneous version than in the sequential version.

Insert Table 8 about here

Aside from the correlations with reading rate, however, our data provide no evidence that the short-term memory manipulation affected the relationship between verbal ability and RT on any of our tasks. Correlations involving simulataneous and sequential conditions are almost identical. Schwartz's hypothesis does not seem to explain why Hogaboam and Pellegrino failed to find correlations with verbal ability.

<u>Evidence for a general "speed of access to long-term memory" factor</u>. Each of the tasks used in this study requires the subject to access semantic or lexical information from long-term memory. Semantic categorization requires that

information be accessed about both pictures and words. Similarily, the semantic verification (FRUIT APPLE) and semantic matching (APPLE PEACH) tasks require extraction of semantic information about word stimuli. Finally, the stimulus matching task requires that lexical, but not semantic, information be accessed. To what extent do these tasks tap a common dimension of individual variation?

To answer this question we first considered the semantic categorization task. This task involved the classification of both words and pictures. We expected that measures based on these two types of stimuli might reflect distinct abilities and thus be only weakly correlated. To our surprise, the correlation between performance on picture and word stimuli was .99! This finding is consistent with Hogaboam and Pellegrino's data. Obviously measures based on picture and word stimuli can be used interchangeably in our statistical analyses.

We next asked whether or not the different measures of memory access formed a unitary dimension. This was done by submitting RTs for positive responses in each of the various information processing tasks to a principal components factor analysis (Harmon, 1967). The variables utilized are listed in the correlation matrix shown in Table 9. In order to make the measures comparable, the NI-PI difference score was replaced by the NI reaction time score. The analysis was actually done twice: once on the first order correlations shown above the diagonal of Table 9, and once on the partial correlations from which the effects of choice RT have been removed. The partial correlations are shown below the diagonal of Table 9. Partial correlations were used to provide a control for the possibility that a common factor would appear that was actually defined by the ability to manipulate the equipment. The results of the two analyses were virtually identical. In each case a single factor was extracted; it accounted for more than 75% of the common variance and had an eigenvalue greater than 5. The eigenvalues of the remaining factors were all substantially less than 1. Table 10 shows the communalities and factor loadings of the information processing measures on the factor, for each analysis. These results indicate that the different memory access tasks tap a single common factor, and this factor is not an apparatus factor. An additional point of interest is that the unrotated factor is very well defined by either the semantic verification (FRUIT APPLE) or the semantic matching (APPLE PEACH) task; both have loadings of about .9 on the common factor.

Insert Tables 9 and 10 about here

In an extension of this analysis we asked whether common factors underlying the reaction time tasks were related to common factors involved in the verbal tasks. A canonical correlation was done to answer this question (Cohen and Cohen, 1979). The first canonical correlation was .69, and was significant at the .02 level. (The canonical analysis was based on the 52 subjects for whom all data was available.) The remaining two canonical correlations were not significant, which is consistent with our previous conclusion that a single factor explains individual performance on the information processing tasks. This result should, however, be treated with caution. Canonical correlations maximize capitalization on chance fluctuations in the data. Thus .69 should be regarded as an extreme upper bound for the correlation between memory access and reading ability in college students.

Is it possible to improve the prediction of verbal performance by using an information processing task that is not heavily dependent upon speed of retrieval of information in permanent memory? This question motivated the inclusion of the sentence-picture comprehension task in the battery of tests. The multiple correlation between the Nelson-Denny reading comprehension score and two independent variables, semantic matching (the best measure of the encoding factor defined in the principal components analysis) and sentence-picture comprehension, was .46, ($p < .001$). The partial correlation between sentence-picture comprehension and reading comprehension, controlling for semantic matching, was .31 ($p < .02$). Thus the sentence-picture comprehension measure accounts for variance in reading ability which is not associated with the encoding measure. This is consistent with similar findings by Palmer et al. (Note 1) in a more extensive analysis of reading comprehension.

Discussion

This experiment was designed to investigate the relationship between verbal ability and speed of access to information in long-term memory. Although our methods were similar to those of Hogaboam and Pellegrino, the design was more powerful both in terms of the number of subjects and the amount of data per subject. Unlike Hogaboam and Pellegrino, we found a positive relationship between verbal ability and several RT measures. However, this experiment was not an exact replication of Hogaboam and Pellegrino's, since we were unable to use exactly the same test of verbal aptitude that they used. Carroll and Maxwell (1979) stated that it was important to resolve the discrepancy between Hogaboam and Pellegrino's results and other studies in the literature. Although of course this cannot be proven, perhaps the discrepancy was due to sampling fluctuations in their study having masked a phenomenon that, in absolute terms, is not a large one.

Hogaboam and Pellegrino's research was motivated by an important consideration. On the basis of results obtained using the stimulus matching paradigm, several investigators have drawn the general conclusion that there is a relationship between (written) verbal skills and speed of access to overlearned information in long-term memory. If the conclusion is sound, then a similar relationship should be found using paradigms other than stimulus matching. In the present study, such a relationship was found, both for the semantic categorization task and for the related semantic verification and semantic matching tasks. Another study in our laboratory replicated this result for semantic verification (Palmer, et al., Note 1). Goldberg, et al (1977) report positive results using a semantic matching task. Our factor analytic results indicate that these paradigms all share a common factor that may be interpreted as speed of access to information in semantic memory. The canonical correlations (and the correlations between the Nelson-Denny measures and markers of the factor) indicate that this factor is related to individual differences in verbal ability.

Performance in a sentence-picture comprehension task also predicted reading comprehension scores, confirming similar observations by Baddeley (1968) and Lansman (1978). Sentence-picture comprehension has a predictive power beyond that provided by a test of retrieval of information from semantic memory. This suggests that there are at least two mechanisms of information processing involved in comprehending verbal material: retrieving the meaning of symbols and manipulating meaning in working memory.

Whether the memory retrieval factor identifiable here is the same as a factor that might be defined by tests of speed of access to information in episodic memory (i.e. whether there is a general memory access factor) is at present an open question. There is tangential evidence that semantic and episodic retrieval factors are not identical. Underwood, Boruch, and Malmi (1978; see also Hunt's (1980) reanalysis of their data) have reported that episodic memory tests are not related to performance on verbal aptitude tests.

The correlations reported in this and other studies are typically in the .3 range. Verbal aptitude is defined by performance on relatively complex tasks and undoubtedly depends upon knowledge as well as upon such aspects of information processing as accessing overlearned information. Other information processing traits, including the ability to hold information in short-term memory and to focus attention, are also involved in tests of verbal ability. It would be naive to expect that any one information processing trait (let alone one paradigm) would provide the single explanation of such a complex ability. It does seem clear that the process of accessing over-learned material is one of the important individual difference variables that underlies skilled verbal performance.

REFERENCE NOTES

1. Palmer, J.C., MacLeod, C. M., Hunt, E., Davidson, J.E. Some relations between information processing and reading. (In preparation.)

REFERENCES

Baddeley, A.D. A 3-minute reasoning test based on grammatical transformation. Psychonomic Science, 1968, 10, 341-342.

Baddeley, A. D. The Psychology of Memory. New York: Basic Books, 1976.

Battig, W.F. & Montague, W.E. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. Journal of Experimental Psychology Monographs, 1969, 80 (3, Part 2).

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

Clark, H.H., & Chase, W.G. On the process of comparing sentences against pictures. Cognitive Psychology, 1972, 3, 472-517.

Cohen, J. & Cohen, P. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences (2nd Ed.) Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1975.

Goldberg, R.A., Schwartz, S. & Stewart, M. Individual differences in cognitive processes. Journal of Educational Psychology, 1977, 69, 9-14.

Harmon, H. Modern Factor Analysis (2nd Ed.) Chicago: University of Chicago Press, 1967.

Hogaboam, T.W. & Pellegrino, J.W. Hunting for individul differences in cognitive processes: Verbal ability and semantic processing of pictures and words. Memory & Cognition, 1978, _6_, 189-193.

Hunt, E. Mechanics of verbal ability. _Psychological Review_, 1978, _85_, 109-130.

Hunt, E. From observations on college students to conclusions about people. In M. Friedman, J.P. Das, & M. O'Connor (Eds.) _Intelligence and Learning_. New York: Plenum, (In press).

Hunt, E., Lunneborg, C. & Lewis, J. What does it mean to be high verbal? _Cognitive Psychology_, 1975, _7_, 194-227.

Kucera, H., & Francis, W.N. _Computational analysis of present-day American English_. Providence: Brown University Press, 1967.

Lansman, M. _Paper and pencil measures of cognitive processes_. Paper presented at the meeting of the Western Psychological Association, Seattle, April, 1977.

Lansman, M. _An attentional approach to individual differences in immediate memory_ (Technical Report). Seattle: University of Washington, Department of Psychology, June 1978.

Matarazzo, J.D. _Wechsler's measurement and appraisal of adult intelligence_. (5th Ed.) Baltimore: Williams & Wickens, 1972.

Nelson, M.J., & Denny, E.C. _The Nelson-Denny reading test_. Boston: Houghton Mifflin Co., 1960.

Posner, M.I. _Chronometric explorations of mind_. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1978.

Posner, M.I. & Mitchell, R.F. Chronometric analysis of classification. _Psychological Review_, 1967, 74, 392-409.

Walsh, K.W. _Neuropsychology: A Clinical Approach_. New York: Church Hill Livingstone, 1978.

Schwartz, S. _Verbal ability and the controlled-automatic_ processing distinction. Paper presented at the NATO International Conference on Intelligence and Learning, York, England, July 1979.

Underwood, B.J., Boruch, R.F., & Malmi, R.A. Composition of episodic memory. Journal of Experimental Psychology: General, 1978, 107, 393-419.

FOOTNOTES

This research was supported by Grant MH-2179 from the National Institute of Mental Health to the University of Washington. Many of the ideas for the design of this study were inspired by discussions with Steven Schwartz. We are grateful to Simon Farr for suggesting the semantic matching paradigm and for his helpful criticism, and to Steven Yantis for his assistance with the data analysis. Reprints may be requested from Earl Hunt, Department of Psychology, NI-25, University of Washington, Seattle, Washington 98195.

Footnote 1. The correlations between the semantic categorization measures and the WPC verbal scores, shown in Table 5, were low and non-significant. It appears that the difference between these correlations and those involving the Nelson-Denny scores is due to fluctuations in small samples; when we recomputed the latter correlations using only those subjects who had WPC scores, they were also low and non-significant.

Table 1

Schedule of Tasks

Session							
1.	Semantic Categorization	Nelson-Denny Vocabulary	Nelson-Denny Reading Comprehension	Paper & Pencil Sentence-Picture Comprehension	Semantic Categorization		
2.	Two-Choice Reaction-Time	Word-Matching Simultaneous Presentation of Stimuli	Word-Matching Sequential Presentation of Stimuli	Semantic Verification Simultaneous Presentation	Semantic Verification Sequential Presentation	Semantic Matching Simultaneous	Semantic Matching Sequential Presentation
3.	Two-Choice Reaction-Time	Semantic Matching Sequential	Semantic-Matching Simultaneous Presentation of Stimuli	Semantic Verification Sequential	Semantic Verification Simultaneous	Word-Matching Sequential	Word-Matching Simultaneous

Table 2

	TIME 1		TIME 2	
	Positive Responses	Negative Responses	Positive Responses	Negative Responses
High freq., words	653 msec	703 msec	622 msec	670 msec
High freq., pictures	637 msec	698 msec	592 msec	672 msec
Low freq., words	711 msec	725 msec	656 msec	694 msec
Low freq., pictures	682 msec	720 msec	609 msec	684 msec

Mean reaction time for correct responses to each stimulus type for each presentation of the Semantic Categorization Task.

Table 3

	SEMANTIC VERIFICATION		SEMANTIC MATCHING	
	Positive Responses	Negative Responses	Positive Responses	Negative Responses
High freq. - SIM	757 msec	862 msec	743 msec	861 msec
Low freq. - SIM	789 msec	867 msec	775 msec	854 msec
High freq. - SEQ	561 msec	638 msec	536 msec	597 msec
Low freq. - SEQ	589 msec	639 msec	557 msec	598 msec

Mean Reaction times for correct responses to each stimulus type for the Semantic Verification (BIRD-ROBIN) and Semantic Matching (ROBIN-SPARROW) tasks. "SIM" indicates simultaneous presentation of the stimuli and "SEQ" indicates sequential presentation.

Table 4

STIMULUS TYPE	SIMULTANEOUS PRESENTATION	SEQUENTIAL PRESENTATION
Name Identical Words (NI)	700 msec	497 msec
Physically Identical Words (PI)	579 msec	478 msec
Different Words in the Same Case	707 msec	542 msec
Different Words in Different Cases	729 msec	544 msec
Subjects	69	69

Mean reaction time for correct responses to each stimulus type for the Stimulus Matching Task

Table 5

Descriptive Statistics for the Psychometric Measures

MEASURE	MEAN # CORRECT	READING RATE PER MINUTE	STANDARD DEVIATION	RANGE	N
elson-Denny Vocabulary	58	—	15	27-98	7
elson-Denny Reading Comprehension	27	—	6	8-36	7
elson-Denny Reading Rate	—	301 Words	95	117-615	7
ntence-Picture Comprehension	44	—	10	21-63	7

Table 6

Correlations Between Processing Speed Measures and Verbal Measures

TIME 1

Processing Speed Measures	Reliabilities	Nelson-Denny Composite	Nelson-Denny # Correct Vocabulary	Nelson-Denny Reading Comprehension	Nelson-Denny Reading Rate	WPC Verbal Composite
High freq., positive resp., words	.95 **	-.21	-.15	-.25*	.07	.01
High freq., negative resp., words	.96 **	-.24*	-.20	-.26*	-.07	.02
High freq., positive resp., pictures	.96 **	-.13	-.08	-.18	.03	.09
High freq., negative resp., pictures	.93 **	-.28*	-.23*	-.29*	-.08	-.09
Low freq., positive resp., words	.94 **	-.29*	-.22	-.32**	.02	.01
Low freq., negative resp., words	.95 **	-.29*	-.25*	-.30*	-.05	-.01
Low freq., positive resp., pictures	.95 **	-.28*	-.20	-.33**	.00	.04
Low freq., negative resp., pictures	.94 **	-.31**	-.25*	-.34**	-.02	.02
Subjects	55	54	54	54	54	26

Table 6 (continued)

Correlations Between Processing Speed Measures and Verbal Measures

TIME 1 plus TIME 2

ng Speed Measures	Reliabilities	Nelson-Denny Composite	Nelson-Denny # Correct Vocabulary	Nelson-Denny Reading Comprehension	Nelson-Denny Reading Rate	WPC Verbal Composite
q., positive resp., words	.98**	-.27*	-.18	-.32**	.03	-.15
q., negative resp., words	.98**	-.28*	-.20	-.33**	-.05	-.17
q., positive resp., pictures	.97**	-.19	-.12	-.25	.03	-.06
q., negative resp., pictures	.97**	-.29*	-.22	-.33**	-.05	-.24
., positive resp., words	.97**	-.29*	-.21	-.33**	.02	-.18
., negative resp., words	.98**	-.27*	-.21	-.30*	-.04	-.16
., positive resp., pictures	.98**	-.28*	-.19	-.36**	.01	-.09
., negative resp., pictures	.97**	-.28*	-.21	-.32**	.00	-.12
Subjects	52	51	51	51	51	25

Table 7

Correlations Between Processing Speed Difference Scores and Verbal Measures

TIME 1

g Speed Measures	Reliabilities	Nelson Denny Composite	Nelson-Denny # Correct Vocabulary	Nelson-Denny Reading Comprehension	Nelson-Denny Reading Rate	WPC Verbal Composite
tures, high freq., pos. resp.	.27 *	-.17	-.17	-.14	.11	-.24
tures, high freq., neg. resp.	-.26 *	.20	.17	.19	.06	.43 *
tures, low freq., pos. resp.	.33 **	.04	-.02	.11	.07	-.13
tures, low freq., neg. resp.	.09	.08	.03	.14	-.06	-.09
freq., pos. resp., words	.31 *	-.31 *	-.28 *	-.30 *	-.11	-.01
freq., neg. resp., words	.06	-.21	-.20	-.18	.07	-.11
freq., pos. resp., pictures	.48 **	-.41 **	-.33 **	-.44 **	-.06	-.09
freq., neg. resp., pictures	.01	-.08	-.04	-.11	.15	.31 *
Subjects	55	54	54	54	54	26

Table 7 (continued)

Correlations Between Processing Speed Difference Scores and Verbal Measures

TIME 1 plus TIME 2

g Speed Measures	Reliabilities	Nelson-Denny Composite	Nelson-Denny # Correct Vocabulary	Nelson-Denny Reading Comprehension	Nelson-Denny Reading Rate	WPC Verbal Composite
tures, high freq., pos. resp.	.19	-.23*	-.21	-.21	-.01	-.31
tures, high freq., neg. resp.	-.16	.13	.14	.10	.03	.46**
tures, low freq., pos. resp.	.37**	-.04	-.12	.07	.06	-.35*
tures, low freq., neg. resp.	.17	-.00	-.05	.06	-.16	-.16
freq., pos. resp., words	.43**	-.20	-.18	-.18	-.02	-.18
freq., neg. resp., words	.05	-.05	-.10	.04	.04	.03
freq., pos. resp., pictures	.35**	-.38**	-.28*	-.45**	-.08	-.12
freq., neg. resp., pictures	.16	.07	.07	.06	.22	.49**
Subjects	52	51	51	51	51	25

Table 8

Information Processing Task	N	NELSON - DENNY COMPONENT SCORE: Composite	Vocabulary	Comprehension	Reading Rate	Reliabilities
Semantic Categorization	52	-.28*	-.20	-.33**	-.02	.99
Semantic Verification-SIM	69	-.24*	-.20*	-.25*	-.09	.97
Semantic Verification-SEQ	69	-.29**	-.27*	-.28*	-.29**	.98
Semantic Matching-SIM	66	-.30**	-.25*	-.31**	-.09	.98
Semantic Matching-SEQ	67	-.29**	-.22*	-.34**	-.20	.98
NI-PI RT; Word Matching-SIM	69	-.25*	-.26*	-.20*	-.01	.61
NI-PI RT; Word Matching-SEQ	69	-.25*	-.23*	-.24*	-.20	.81

Correlation between Nelson-Denny Component Scores and various Reaction Time measures for positive response. "SEQ" indicates sequential presentation, "SIM" indicates simultaneous presentation. Semantic Categorization is the replication of Hogaboam and Pellegrino's results.

3

Table 9

Positive Responses Information Processing Tasks	1	2	3	4	5	6	7
1. Semantic Categorization		.58	.53	.50	.47	.58	.5
2. Semantic Verification - SIM	.39		.80	.92	.77	.79	.7
3. Semantic Verification - SEQ	.33	.67		.76	.87	.69	.7
4. Semantic Matching - SIM	.34	.81	.71		.79	.74	.7
5. Semantic Matching - SEQ	.37	.73	.85	.85		.65	.8
6. NI, Word Matching - SIM	.43	.67	.45	.56	.46		.7
7. NI, Word Matching - SEQ	.42	.59	.74	.69	.74	.67	

Correlations used in the principal components analysis. The upper right diagonal shows the raw correlations. The lower left diagonal shows the correlations after choice reaction-time has been partialled out.

Table 10

Positive Responses Information Processing Variables	First Order Analysis		Partial Correlation Analysis	
	Communality	Factor Loading	Communality	Factor Loading
;ategory Verification	.39	.60	.38	.59
;emantic Verification-SIM	.89	.93	.85	.90
;emantic Verification-SEQ	.81	.89	.80	.89
;emantic Matching-SIM	.87	.89	.84	.89
emantic Matching-SEQ	.83	.87	.87	.88
I, Word Matching-SIM	.75	.84	.79	.79
I, Word Matching-SEQ	.79	.88	.78	.85

Communalities and factor loadings for information processing variables, based on analysis of original ('First Order') correlations and on partial correlations removing variance associated with choice reaction-time.