DOCUMENT RESUME

ED 171 736 TM 008 742

AUTHOR	Allen, Ted W.; And Others
TITLE	An Information Processing Approach to Performance Assessment: III. An Elaboration and Refinement of an Information Processing Performance Battery. Technical Report No. 3, November 16, 1977 through November 30, 1978.
INSTITUTION	American Institutes for Research in the Behavioral Sciences, Washington, D.C.
SPONS AGENCY	Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
REPORT NO	AIR-58500-TR
PUB DATE	30 Nov 78
CONTRACT	N00014-76-C-0871
NOTE	129p.; For related documents, see TM 008 762, TM 009 193-194; Table 6 may be marginally legible
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	Aptitude Tests; Cognitive Measurement; *Cognitive Processes; *Cognitive Tests; Higher Education; *Individual Differences; Learning Theories; Memory; *Task Analysis; Test Construction; Testing; Test Reliability; *Test Validity; Verbal Tests
IDENTIFIERS	*Information Processing Performance Battery

ABSTRACT

An information processing approach was applied to the development and validation of a test battery intended for personnel selection, classification, and guidance; and design of training programs. The approach specifies that tests should measure specific cognitive processes and basic abilities, rather than prior experience. Tests should be short, easy to administer, statistically independent, and reliable. Taken as a whole, the test should not be designed to predict ability or interest in a particular job. Sixty eight college students were tested twice on eight tasks: (1) physical match; (2) letter rotation; (3) scan and search; (4) set membership; (5) letter recall; (6) mental addition; (7) sentence recall; and (8) sentence recognition. Results showed alternate forms of the tasks were comparable with previous findings for all but task 8. The adequacy of the tasks as measures of individual differences and as measures of information processing operations varied from task to task. The first four tasks were selected for the battery; tasks 5 and 6 were considered marginal (needed improvement); and tasks 7 and 8 were excluded. (Sample items from each task are appended.) (CP)

An Information Processing Approacn to Performance Assessment:

III. An Elaboration and Refinement of an Information Processing Performance Battery

Ted W. Allen
Andrew M. Rose
Leslie J. Kramer

Prepared for the Personnel and Training Research Programs, Psychological Sciences Division, Office of Naval Research, Arlington, Virginia.

Contract No. N00014-76-C-0871

TECHNICAL REPORT
November 1978

AMERICAN INSTITUTES FOR RESEARCH / 1055 Thomas Jefferson Street, NW, Washington, DC 20007

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report No. 3

2. GOVT ACCESSION NO.:

3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): "An Information Processing Approach to Performance Assessment: III. An Elaboration and Refinement of an Information Processing Performance Battery"

5. TYPE OF REPORT & PERIOD COVERED: Technical Report 16 Nov. 1977 - 30 Nov. 1978

6. PERFORMING ORG. REPORT NUMBER: AIR 58500-TR

7. AUTHOR(s):
T. W. Allen
A. M. Rose
L. Kramer

8. CONTRACT OR GRANT NUMBER(s): N00014-76-C-0871

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
American Institutes for Research
1055 Thomas Jefferson St., N.W.
Washington, D.C. 20007

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61153N; RR042-04-01; RR042-04; NR 150-391

11. CONTROLLING OFFICE NAME AND ADDRESS:
Personnel and Training Research Programs,
Office of Naval Research - Code 458
Arlington, Virginia 22217

12. REPORT DATE: November 30, 1978

13. NUMBER OF PAGES: 125

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):

15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

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

18. SUPPLEMENTARY NOTES

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

Personnel Technology
Individual Differences
Personnel Assessment
Information Processing
Cognitive Processing

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

This report describes the third study in a program of research regarding the development and validation of a comprehensive standardized test battery that can be used as an assessment device for the evaluation of performance in a wide variety of situations. The standardized battery is being designed to possess high reliability and predictive

20. Abstract (Cont'd.)
validity for a wide variety of criterion tasks. Equally important, the battery is being designed to include tests that possess construct validity: there will be a firm theoretical and empirical base for inferring the information processing structures that the tests purport to measure.

The major purpose of the present study was to determine the properties of a set of tasks selected for inclusion in the test battery. Three questions were of primary interest: the replicability of previous findings with alternate forms of the tasks, the adequacy of the tasks to provide measures of individual differences, and the adequacy of the tasks to provide measures of information processing operations.

Sixty-eight subjects were tested twice on each task. The tasks investigated included:

Physical Match
Letter Rotation
Scan and Search
Set Membership
Letter Recall
Mental Addition
Sentence Recall
Sentence Recognition

In general, the results showed the forms of tasks used here to be quite compatible with previous findings for all tasks but one, Sentence Recognition. Even for this task, there is support in the experimental literature for the obtained findings. The support for individual difference measures and measures of information-processing operations varied from task to task. The summary presents our recommendations regarding the inclusion of various tasks and measures in the battery.

ACKNOWLEDGMENTS

The authors wish to acknowledge the contribution of Dr. Tetsuro Motoyama of American Institutes for Research, who authored the data scoring programs and guided us in the selection and derivation of appropriate statistical analyses.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The Information Processing Performance Battery (IPPB) is an attempt to apply the theories and methods of laboratory-based studies of human cognitive performance to the area of performance assessment. The primary rationale developed through an ongoing series of experiments (Rose, 1974; Rose and Fernandes, 1977; Fernandes and Rose, 1978) is that individuals can potentially be characterized in terms of parameters derived from models of selected information processing tasks. If these parameters can be demonstrated to meet standard test-item criteria, then a test battery comprised of such measures would not only be potentially predictive of performance on a wide variety of real-world tasks but would also be firmly based in theory. Such a test battery would represent a significant advance over standard selection and placement tests; it would promote increased understanding of the cognitive operations involved in any criterion task shown to be related to constructs in the test battery.

The basic approach involved in this research is exemplified by Rose (1974). The strategy was to select experimental tasks from the psychological literature that had been demonstrated to be valid measures of information processing constructs. A set of tasks was selected; each task was adapted to fit logistic demands of time and equipment, and then administered to a large group of subjects. Extensive correlational analyses were conducted to determine the relationships among the tasks and the individual task reliabilities. These procedures resulted in a set of tasks which were reliable, statistically independent, and (arguably) possessing high construct validity.

The Rose and Fernandes (1977) study extended this approach by hypothesizing a set of constructs ("operations") which were used to model performance for each task. Since most of the task parameters could be cast as time measures, it was possible to employ regression techniques to "converge" upon these operations. Some fairly simple assumptions, such as linear additive stages and independence of operations, led to the estimation of durations for some of these operations. More importantly, the generation of these constructs provided a valuable heuristic device for the interpretation of task performance and provided an initial empirical basis for the isolation of basic information processing components.

Fernandes and Rose (1978) attempted to extend the methodology to the realm of memorial tasks. Based on a study by Underwood, Boruch, and Malmi (1977), several memory-related tasks were selected for more detailed evaluation. Although modeling of these tasks in terms of operations was not attempted, it was possible to examine the obtained relationships among the tasks for commonalities that could be interpreted in information processing terms.

The present study is similar to the above studies in that the general approach was the same: the literature was reviewed in order to select candidate paradigms, these paradigms were adapted to meet logistic limitations, and the tasks were administered twice to subjects. The major differences between this study and the others are that first, several tasks previously included in the IPPB were readministered, primarily to test for alternate-form consistencies and to capitalize on the previous findings for interpretation of results. Second, a number of "new" treatments were built into this study. These treatments extended the theoretical underpinnings of the selected tasks, thus allowing for "stronger" interpretations of the phenomena under study. Third, this study made greater use of analysis of variance techniques for the isolation of potential individual difference parameters. These changes in analytical procedures are discussed below.

General Task and Operations Overview

The task selection process was based upon an extension of the approach advocated by Rose and Fernandes (1977). They suggest that information processing tasks can be described as requiring some combination of information manipulation operations, where an operation represents a set of control processes that must be performed in order to successfully complete the task (see Carroll, 1974). Rose and Fernandes arrived at eight specific operations necessary to describe the tasks used in their study. In the present study we have refined their list of operations, limiting the scope of some operations, more clearly delineating the boundaries of others, and adding new operations in order to cover a wider range. These operations are described below.

Encoding: The operation by which information is input into the information processing system. The encoding processes are primarily concerned with making the preliminary analysis of external information necessary to get that information into the system. Some processes that might occur in the context of this operation are: attribute or feature extraction, selection, erasure, masking, and temporal summation.

Abstraction: The operation by which new structures are generated from information already in the system. The abstraction processes are primarily concerned with generation of higher level structures beyond those initially encoded in the system. Processes that might occur in this operation include: Neisser's (1967) synthesis, schema formation, concept formation, and categorization.

Transformation: The operation by which information is converted into an equivalent structure that is required for successful task performance. In contrast to abstraction, the processes involved in transformation do not involve the generation of new (unencoded) information but rather require the application of some stored rules to the information structure already present. Some processes that could occur in a transformation operation are: rotation, chunking, chaining, updating, and segmentation.

Recoding: The operation by which the information is converted into a corresponding structure of different form or composition. In contrast to transformation, the recoding processes require the generation of new information that is not an end-product of the encoding operation. In contrast to abstraction, the recoding operation does not require a "synthesis" or "formation" so much as a substitution of corresponding structures of different types. Recoding processes would include, for example, changing the structure of information in the system from a verbal form to an imaginal form, or from acoustic to semantic.

Storage: The operation by which new information is integrated into the existing organization of information structures and the processes impinging on the information while there. Storage processes might include: decay, degradation, displacement, tagging, unlearning, and Estes' (1972) reverberation.

Retrieval: The operation by which previously stored information is made available to the processing system. Whether it is indicative of the extent of research on the operation or the primitive nature of the operation, the literature contains no exemplary processes for the retrieval operation.

Searching: The operation by which an information structure is examined for the presence or absence of one or more properties. The information structure examined may be one already in the processing system or one external to it (e.g., a visual array). The searching operation might include processes such as attribute verification and scanning.

Comparison: The operation by which two entities, either internal or external to the processing system, are judged to be the same or different. An example of the processes involved in the comparison operation would be a matching process.

Decision and Response: The operation by which the appropriate motor action is selected and executed. Among the many processes in this operation are monitoring and evaluation processes.

The operations described above represent only a part of the information necessary to completely describe task performance. At the very least, one might need information about the contents or form of the information structure to which the operation is applied (e.g., a comparison operation might respond differently to pictorial, acoustic or semantic information) and one needs information about the state of system stress (i.e., the amount of incoming information and required speed of processing). For the tasks described below we shall assume that content of the operations is irrelevant. In most cases the design of the tasks and data comparisons support this assumption; however, since we cannot at this time precisely specify the effects of contents or the interaction of contents and operations we have a problem of "identifiability" and would have to make this or an equivalent assumption regardless of task and data comparison considerations. System stress can also be assumed to be comparable in the different tasks. The possible effects of these two variables will be entertained where relevant in the discussion.

Listed below is a description of the eight tasks included in the present study. Each description is accompanied by a discussion of the operations involved. The nature of their involvement in each task is summarized in Table 1.

Task Descriptions

Letter Recall Task. The letter recall task is a serial, short-term retention task designed to embody the spirit of both the memory span task and Waugh and Norman's (1965) probe-digit task. These two tasks are used to investigate the limited (storage or processing) capacity of the memory system. In a typical memory span task, subjects study lists of items of

TABLE 1.
Operational Analysis of Tasks

Tasks	Information Processing Operations								
	Encoding	Abstraction	Transformation	Recoding	Storage	Retrieval	Search	Comparison	Decision-response
1. Letter recall	minor	minor	minor	minor	**major**	**major**	minor	minor	minor
2. Mental addition	minor	minor	**major**	minor	**major**	**major**	minor	minor	minor
3. Sentence recall	minor	**major**	minor	minor	**major**	**major**	minor	minor	minor
4. Sentence recognition	minor	**major**	minor	minor	**major**	minor	**major**	minor	minor
5. Letter rotation	minor	minor	**major**	minor	minor	minor	minor	**major**	minor
6. Physical match	minor	minor	minor	minor	minor	minor	minor	**major**	minor
7. Set membership	minor	minor	minor	minor	**major**	minor	**major**	minor	minor
8. Scan and search	minor*	minor	minor	minor	**major**	minor	**major**	minor	minor

major = operation is of MAJOR importance in determining task performance.
minor = operation is of MINOR importance in determining task performance.
* For degraded stimuli, encoding is of MAJOR importance

increasing length which must be recalled in order from a single presentation. Typical results show the percentage of items recalled to remain at about 100% for the first 4 ± 1 items and to decrease rapidly thereafter. Usually the memory span is defined as the number of items recalled correctly 50% of the time and is typically 7 ± 2 items. A better measure of capacity however might be the amount recalled correctly at or near 100% of the time in which case 4 ± 1 would be the capacity estimate.

The probe-digit task consists of the presentation of a long series of items (e.g., 14 digits) and the immediate probing of retention by re-presenting one of the items from the series. The subject is to recall the item which immediately followed the probe in the series. The task typically demonstrates the displacement effect, that later items in the series tend to displace or knock out earlier items from the limited capacity system.

In the present task the subject studies lists of varying length (5-10 letters) and must recall the last five letters in order. Since the subject does not know when the list will end and is instructed to retain only the last five items, we would expect, theoretically at least, the subject to be engaged in a procedure of filling the system to near capacity (4 ± 1 items) and then displacing the oldest items as new items are processed. Variance in task performance should result from two sources: (1) the size of the individual's memory capacity and (2) the individual's ability to control the displacement process. Thus with respect to the operations discussed previously we expect the storage operations and retrieval operations to be the major factors in task performance.

Procedure. In this task subjects recalled the last five letters of a series of letters. In each trial subjects heard a series of letters presented at the rate of one letter per second. At the end of the series subjects heard the word "RECALL" and were given fifteen seconds to write down the last five letters of the series in the order in which they occurred. Instructions told subjects to leave a blank space if they couldn't remember a letter, and if they could remember the letters but not their order, they were to guess the positions of the letters.

The following shows the events of two trials:

STIMULI	RESPONSES
"K C M T N, Recall"	K C M T N
"A M G S K P F C, Recall"	S K P F C

Stimuli and design. This task was presented twice a day. Each administration consisted of five trials explicitly designated as practice and twenty-two critical trials with no feedback provided. There were three trials for each series of 5, 6, 7, 8, 9, and 10 letters. These 18 trials were arranged randomly with respect to the number of letters in each trial with the restrictions that the first trial be a five-letter trial and no consecutive trials have stimuli of the same length. The remaining four trials were non-scored five-letter series. These occurred as the second, fourth, sixth, and twenty-second trials and were inserted with the purpose of discouraging subjects from ignoring the first few letters of any trial.

The stimulus for each trial was constructed by random selection from the set of consonants, excluding W and Y, with the constraints that no letter be repeated during a trial and a maximum of three acoustically similar letters be permitted to occur consecutively within a trial.

Variables. Three principal measures were computed for the task. The order-recall measure consisted of the number of letters recalled from the last five letters in the series in the correct positions; thus credit was given only if a letter recalled was one of the last five letters in the series and if

that letter was recalled in its correct position. The derived-free-recall consisted of the number of letters recalled from the last five letters in the series irrespective of the correspondence of their positions. Here credit was given for any letter recalled from the last five letters, regardless of the position in which it was recalled. Lastly, the series-recall measure consisted of the number of letters recalled from any position in the series. For each principal measure, two derived measures were obtained: mean performance and the slope of the function relating amount recalled and condition.

Mental Addition Task. The mental addition task is a replication of a task designed by Hitch (1978) to study processing capacity in the memory system. Hitch read to his subjects two numbers (e.g., "four hundred eighty-seven plus twenty-six") which they had to add together mentally, without writing or seeing the addends, and report the sum. Separately he varied:

(1) the size of the addends (i.e., two three-digit numbers or a three-digit number and a two-digit number),

(2) the amount and location of "carrying" operations (i.e., no carrying, carrying into tens position, carrying into hundreds position, and carrying into both tens and hundreds positions).

Hitch found that the time to complete the addition was greater for the two three-digit addends than for the three-digit and two-digit addends, and that the time increased with the number of carrying operations required.

The mental addition task in the present study combines the two above-mentioned variables. Theoretically, variance in task performance across individuals should depend on three factors:

(1) The individual's ability to store and maintain the addends and carries,

(2) The individual's ability to rapidly retrieve both the rules of addition and the addends and carries,

(3) The individual's ability to actually perform the addition; in our terminology, this can be categorized as a transformation operation.

Thus, the major operations involved in task performance should be the storage, retrieval, and transformation operations.

Procedure. Subjects were required to mentally add two multiple-digit numbers. Subjects heard a cue ("READY") followed by two numbers. They then had five seconds to add the numbers before they heard the word "STOP". Subjects were instructed to write down each digit of the answer as soon as the addition for that digit was complete. They were also told not to guess. Here is the sequence of events of two trials:

STIMULI	RESPONSES
"Ready, two hundred fifty-eight plus three hundred twenty-seven, ... (five second pause) ..., STOP"	5 8 5
"Ready, four hundred ninety-two plus two hundred forty-one, ... (five second pause) ..., STOP"	7 3 3

Stimuli and design. This task was administered twice each day and subjects completed twenty-four trials during each administration. They received no feedback. Fifty percent of the trials required the addition of two 3-digit numbers; in the remaining trials subjects added a 2-digit number to a 3-digit number. The trials were divided equally with respect to the number of carry-overs required to complete the addition so that 25% of the trials required no carries, 25% required a carry from the one's column to the ten's column, 25% required a carry from the ten's column to the hundred's column, and 25% required carries from both the one's and the ten's columns to the ten's and the hundred's columns respectively. The answers to the problems were chosen at random (excluding zero), and each answer was arbitrarily assigned to one of the four categories of carry-over. Then the addends were chosen so that the problem would correspond to the category of carry-over to which it was assigned.

The digits of the addends were modified to provide approximately equal representation of each digit 1-9 in each position of the sum, with the following exceptions: 1) in sums with two 3-digit addends, the number "9" is not found in the hundred's place of the first or second addends and large numbers are underrepresented in those two positions; 2) in sums of a 3-digit addend plus a 2-digit addend, the number "9" is not found in the hundred's place (of the first addend) and large numbers are underrepresented in this position.

Variables. The main data were the number of uncomputed (blank) positions in the reported sum and the number of correct positions in the reported sum. Mean performance for these measures was obtained for each of the conditions.

Sentence Recall Task. The sentence recall task is an original task designed for this study. It embodies much of the theoretical essence of a free-recall cluster analysis in studying the individual's attempt to organize incoming information. In the task, subjects are read a series of simple sentences, each sentence expressing an idea or attribute relating to one of four topics. At the end of the series subjects are to contiguously recall the ideas associated with each topic. In the present design, the number of ideas associated with each topic is varied (4, 5, or 6).

The task is designed to get subjects to integrate the information presented in the simple sentences on each topic. To the extent that individuals performed such integration, they should be able to decrease the memory load requirements of the task. The major operations determining performance then should be: abstraction (organization and integration of sentence information), storage, and retrieval.

Procedure. In this task subjects had to recall and organize information. Subjects heard a list of sentences on four alternating topics. At the end of the list they heard a cue ("RECALL") and had approximately two minutes to recall the information in the sentences and organize it by topic. Subjects were told that they could either organize the information into one complex sentence or use phrases to describe the information they had heard.

The following shows the sequence of events for part of a trial:

STIMULI
"The party was loud."
"The snake was in the cage."
"The party was a surprise."
"The cage was glass."
"The party was a success."
"The snake shook its rattles."
"RECALL"

RESPONSES
The loud surprise party was a success.
OR
Party was a surprise, loud, a success.
The snake in the glass cage shook its rattles.
OR
Snake in cage shook rattles, cage was glass.

Stimuli and design. Subjects completed one practice and three critical trials of this task each day. They did not receive feedback after the practice. The acquisition list for each trial was based on four topics. Each sentence in the acquisition list explained a relationship between attributes of one of the four topics. Each sentence had one subject and one word or phrase describing the subject or an action by the subject. The sentences were presented alternately by topic so that every fourth sentence was on the same topic. In the first trial 16 sentences were presented, four for each topic;

in the second trial 20 sentences were presented, five for each topic; and in the third trial 24 sentences were presented, six for each topic.

Variables. Two principal recall measures were obtained for the task. A cluster recall measure consisted of the number of attributes recalled contiguously for each topic. This measure reflects the number of attributes that the subject correctly assigns to a given topic. The item recall measures consisted of the number of nouns, adjectives, and verbs (with the exception of *TO BE*) that were recalled from a topic description. Credit was given in both measures for paraphrases and synonyms.

Sentence Recognition Task. The sentence recognition task is a replication of a task designed by Bransford and Franks (1971) to investigate the acquisition and retention of linguistic ideas. On acquisition in their study subjects were presented sentences containing various combinations of ideas from a complex, four-idea sentence. Retention was tested with a recognition test using only new sentences with ideas from the same complex sentence. Bransford and Franks found that the more ideas in a (new) test sentence the more likely the subjects were to judge that sentence as old with high confidence.

The present task added a modification to the Bransford and Franks design. Where Bransford and Franks used the same syntax in all the derived sentences as in the complex sentence, we have varied whether the syntax is the same or different in the sentences. This modification is intended to show the extent to which the effect is dependent on repetition of syntax.

Performance on the sentence recognition task presumably depends upon the individual integrating and organizing the ideas from the various sentences, such that the individual can no longer distinguish the idea combinations that were presented from those that were the product of integration processes. If syntas does not completely explain the effect, then whatever variance remains should be due to abstraction operations (idea integration), storage operations (retaining sentences in memory) and search operations. Search operations are involved as opposed to retrieval operations because the test is recognition.

Procedure. In this task, subjects decided whether or not they had heard particular sentences from a list of sentences previously presented. An acquisition list of 12 sentences based on two topics was orally presented. Fourteen recognition sentences about the two topics were then presented. After each recognition sentence was read, subjects decided whether or not it was a member of the acquisition list and gave a confidence rating to indicate how certain he/she was that his/her answer was correct. Subjects were told that the wording of a sentence from the recognition list had to be identical to the wording of the corresponding sentence from the acquisition list in order for the sentences to be considered the same. A practice task was completed before each administration of the critical task, but no feedback was provided.

The following shows the sequence of events of part of a task.

ACQUISITION LIST

"The rock rolled down the mountain crushing the hut at the edge of the woods."

"The actress looked at her reflection."

"The rock that rolled down the mountain crushed the hut."

"The actress wearing a hat looked in the mirror."

RECOGNITION LIST	RESPONSE		CONFIDENCE RATING
"The actress looked at her reflection."	(Yes)	No	1
"A hut was at the edge of the woods."	Yes	(No)	2
"The rock that rolled down the mountain crushed the hut."	(Yes)	No	3
"The mirror was cracked."	Yes	(No)	4

Stimuli and design. Subjects completed this task twice each day. Each task was based on two primary complex sentences about different topics. Each primary sentence had two subjects. In one of the primary sentences, each subject had two modifying phrases; in the other primary sentence one subject had three modifying phrases and the other subject had one modifying phrase. For example, the primary sentence "The teenager who lives next door broke the large window on the porch" has two subjects and each subject has two modifying phrases.

The teenager → lives next door.
The teenager → broke the window.

The window → was large.
The window → was on the porch.

In each of the remaining sentences in the task, one or more of the modifying phrases which were found in the primary sentence was deleted. For example, some sentences generated from the above primary sentence are:

The teenager broke the window.
The large window was on the porch.
The teenager who lives next door broke the large window.

In each task, the sentences generated from one of the primary sentences used the same wording as the primary sentence. This was called a consistent set of sentences. The example above is a set of consistent sentences. The sentences generated from the other primary sentence in the task were inconsistent. For example, the following is a primary sentence with a set of inconsistent generatives:

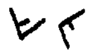
PRIMARY SENTENCE	GENERATIVES
The executive who drank the hot coffee at his desk had his briefcase.	The executive was sitting at his desk.
	At his desk, the executive was drinking hot coffee.
	The coffee was hot and was drunk by the executive with the briefcase.

Of the 14 sentences in the recognition list, two were the two primary sentences, six were sentences from the acquisition list ("OLD") and six were sentences the subjects had not heard before ("NEW"). Thus, both the acquisition list and the recognition list contained twelve non-primary sentences, six on each of two topics; of the six, half were "OLD" and half were "NEW". Of the six generatives of a particular topic, two were made by deleting one modifying phrase from the primary sentence, two were made by deleting two modifying phrases from the primary sentence, and two were made by deleting three modifying phrases from the primary sentence. In the acquisition list, the topics of the twelve sentences were alternated, that is, every other sentence was a generative of the same primary sentence and no generative was based on the same primary sentence as the previous generative. In the recognition list, the generatives plus the two primary sentences were presented in random order.

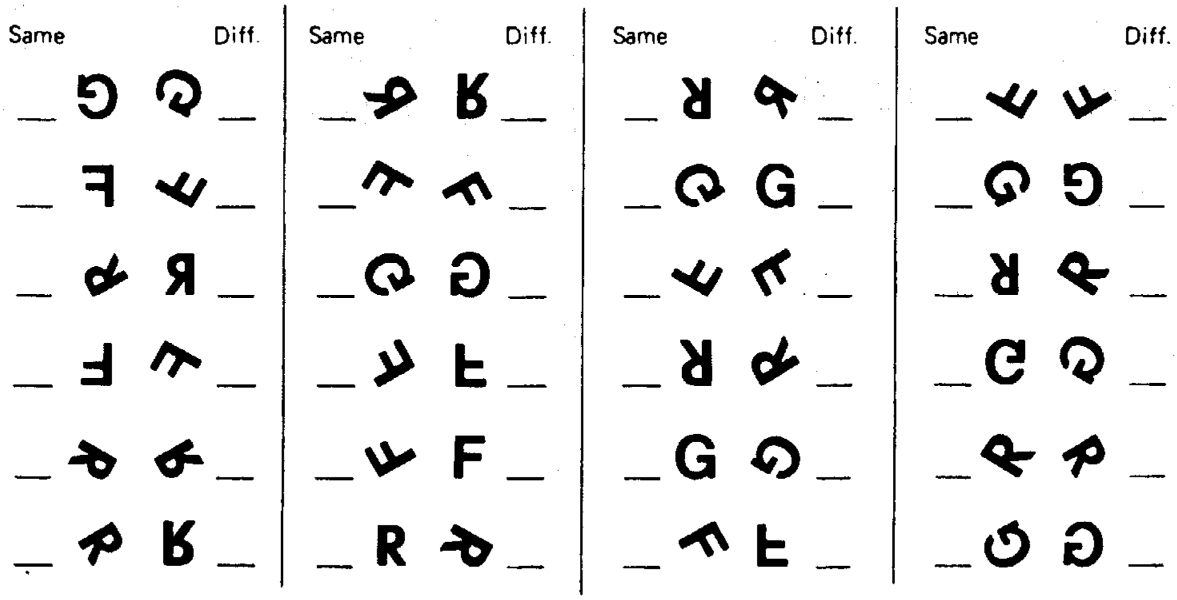
Variables. In this task the data were the yes-no judgments and the ratings of those judgments. The information was combined in a modified rating scale by giving all ratings for the subject's "old" responses a "plus" and all ratings for "new" responses a "minus". Thus, a 10-point rating scale was developed ranging from minus five to plus five (excluding zero). Using this scale, the best-fitting linear functions relating rating and sentence complexity were derived for old and new sentences. Mean number of errors and the overall frequency distribution of the ratings were also obtained.

Letter Rotation Task. This task replicates a task investigated extensively by Lynn Cooper and Roger Shepard (1973). The subject is shown two items and is asked to determine whether they are different objects or the same object from a different perspective. It has been found that the time required to answer this question is a linear function of the degree of rotation required to bring the figures to a common orientation.

With respect to the operations discussed earlier, this task would involve the transformation and comparison operations. The rotation of the objects would involve processes subsumed under transformations. Once the objects have been rotated the comparison operation is required to determine if the objects are the same or different.

Procedure. In this task, subjects decided if two letters were the same or different using two rules. According to Rule 1, two letters were the same if a subject could slide one letter on top of the other and they matched exactly. For example, using Rule 1  is "SAME" and  is "DIFFERENT". Subjects completed the first page of the task using only Rule 1. Subjects were instructed to use Rule 2 for pages 2, 3, and 4. According to Rule 2, two letters were the same if, after mentally rotating one letter, a subject could slide one letter on top of the other and they matched exactly. For example, using Rule 2  and  are "SAME" and  and  are "DIFFERENT". Subjects were given thirty seconds to work on the first page of this task and sixty seconds to work on the second, third, and fourth pages of this task.

The following are examples of trials used in this task:



Stimuli and design. This task was administered once each day. During each administration subjects completed four pages. Each page had four columns of 18 trials each. The stimuli were pairs of capital F's, G's, and R's. In half of the pairs of letters, the second letter was a mirror image of the first, making it impossible to slide the letters on top of one another and have them match. These pairs were "DIFFERENT". In the rest of the pairs of letters one letter was not a mirror image of the other, so they were "SAME". In the second, third, and fourth pages of the task the second letter was rotated with respect to the first letter. On the second page, the second letter of each pair had a 60° rotation with respect to the first letter, on the third page the second letter had a 120° rotation with respect to the first letter. And on the fourth page the second letter had a 180° rotation with respect to the first letter. For example, on the third page, if the first letter was oriented at an angle of 60° clockwise with respect to the vertical axis, the second letter would be oriented at 180° with respect to the vertical axis. The left member of a pair was in one of six positions clockwise with respect to vertical: 0° (vertical), 60°, 120°, 180°, 240°, or 300°. Since three letters were used, each item was one of 72 possible combinations for a particular page. The items were distributed randomly within a page, with the restriction that in each column 50% of the items were "SAME" and 50% were "DIFFERENT".

Variables. The data consisted of the mean time per item for the same - different judgments across the varying degrees of rotation. The percentage of errors in each condition was also computed.

Physical Match Task. The physical match task is a "same-different" classification task derived from an experimental paradigm developed by Posner and Mitchell (1967). The subject is presented two uppercase letters and asked to respond "same" if the two letters are physically identical and "different" otherwise. In terms of our operations, this task represents an uncomplicated assessment of the comparison process.

Procedure. Subjects saw two letters side by side and classified the pair as "SAME" or "DIFFERENT". A pair of letters was to be classified as "SAME" if the items in the pair were the same letter and "DIFFERENT" if the items were not the same letter. Subjects completed a non-timed practice task and the were given 45 seconds to work on the critical task.

The following shows the events of 7 trials:

STIMULI

Same			Diff.
___	E	E	___
___	T	T	___
___	E	H	___
___	H	H	___
___	T	E	___
___	A	A	___
___	E	A	___

RESPONSES

Same			Diff.
√	E	E	___
√	T	T	___
___	E	H	√
√	H	H	___
___	T	E	√
√	A	A	___
___	E	A	√

Stimuli and design. This task occurred once during each session (day). The stimuli were combinations of the letters A, E, H, and T. The task was constructed so that half of the stimuli had a correct response of "SAME" and half had a correct response of "DIFFERENT". The various letter combinations occurred randomly with the restriction that a particular answer not occur more than four times consecutively.

Variables. The data consisted of the mean time per for the same - different judgments. The percentage of errors was also obtained.

Set Membership Task. The set membership task replicates an experimental paradigm developed by Sternberg (1967, 1969) to study information retention. Subjects memorized a sub-span list of items (1 - 4 letters) and had to respond to a series of probes as to whether the probe was in the memory set. Subjects made "yes" or "no" responses to each item. Time to process and respond to the probe increased linearly as the size of the memory set increased.

This task involves two types of operations:

(1) Storage operations are required to maintain the memory set in memory.

(2) Search operations are required to scan the target set for the probe.

Procedure. In this task, subjects decided if letters presented in printed columns were members of a target set. Subjects were instructed to put a check to the left of a stimulus letter (in a column labeled "YES") if the letter was a member of a target set and to put a check to the right of a stimulus letter (in a column labeled "NO") if the letter was not a member of the target set. Subjects were verbally presented with a target set of 1, 2, 3, or 4 letters and were then given thirty seconds to work on the task. After thirty seconds subjects were told to stop working and to write the target letter(s) at the bottom of the page.

The following shows the sequence of events for eight items of a task:

TARGET	STIMULUS Yes		STIMULUS No	RESPONSE Yes		RESPONSE No
"B P Q N"	___	L	___	___	L	✓
	___	N	___	✓	N	___
	___	P	___	✓	P	___
	___	M	___	___	M	✓
	___	N	___	✓	N	___
	___	Q	___	✓	Q	___
	___	V	___	___	V	✓
	___	C	___	___	C	✓

Stimuli and design. Subjects completed two practice and 12 critical trials of this task each day. They received no feedback on their performance. Both the stimuli and the targets were randomly chosen from the set of 19 letters exclusive of vowels, Y and W. Target sets of one, two, three, and four letters occurred randomly with the restriction that no two consecutive targets had the same number of letters. Half of the letters in each stimulus list were members of the target set for that list, and the order in which the letters occurred was random.

Variables. The data consisted of the mean time to process each probe. These times were plotted against the number of elements in the target set and the best-fitting linear function derived. The percentage of errors was also computed.

Scan and Search Task. The scan and search task was originally developed by Ulric Neisser (1967) to assess the memory search process. Subjects search a list of letters for target letters. The list contains a column of letters with four to six letters to a row. Here the subject searches down the column and responds only when a target letter is found, as opposed to the set membership task where the subject responds to every item. Neisser found that search time increased the further down the column the target letter was positioned. Also, and most important to us, he found that with extensive practice, his subjects could search for any one of ten targets simultaneously with the same speed as they could search for one target. In other words, the size of the target set did not affect search times.

In the present version of the task, subjects memorize a sub-span target set (1, 2, or 4 letters) and then search a list of letters for target letters. Performance is tested on "clear" and "degraded" lists. Individual performance should depend upon storage operations to hold the target set in memory and search operations to scan the column for the target. A comparison between the "clear" and "degraded" conditions should reveal effects of the encoding operation since these processes would be required to "clean up" the degraded stimuli.

Procedure. In this timed task, subjects searched columns of groups of letters looking for particular letters. At the beginning of each trial, a target set of one, two, or four letters was verbally presented. Subjects then searched for items which contained one of the target letters, placing a check next to those that did. Subjects received 20 seconds for the one-letter target condition and 30 seconds for the two and four-letter target conditions. At the end of each trial, each subject drew a line under the last item he had looked at and wrote down the letters he had been looking for.

The following are events of part of a typical trial:

TARGET	STIMULUS	RESPONSE
B, F	A U M F Z	A U M F Z ✓
	L N D Q G	L N D Q G
	K L J D P	K L J D P
	Y K T M F	Y K T M F ✓
	X V D B Y	X V D B Y ✓

Stimuli and design. This task was completed twice each day. Subjects received two practice and six critical trials during each administration. No feedback was provided. Half of the stimulus pages were degraded by printing them through a series of filters. The filters were Xerox No. 8R535 Document Carrier and Screens, and were made of a regular pattern of small white dots. During an administration, subjects completed the non-degraded one-letter, two-letter, and four-letter target conditions, in that order, and then completed the degraded targets in the same order.

Each item was a string of five capital letters. There were four columns of 16 items each on each page. Fifty percent of the items in each column contained a target letter and no item contained more than one target letter.

Variables. The principal measure was the mean time to scan each item and make the yes-no judgments. These times were plotted against the number of items in the search set and the best-fitting linear function derived. Separate functions were derived for the clear and degraded conditions. The percentage of errors for each condition was also computed.

METHOD

Subjects

Subjects were 68 female and male students who answered an advertisement in the school newspaper at Georgetown University. They were paid for their participation in the study and were tested in groups ranging from three to eight people.

General Procedure

The subjects participated in two testing sessions, each approximately three hours in duration and scheduled a day apart. All eight tasks were presented in the same order each session. Where appropriate, different stimuli were used in each testing session and the order of stimuli in the task was randomized.

At the beginning of the testing sessions subjects received a booklet containing the instructions and response sheets for all the tasks. Appendix A presents instructions for each task and selected stimuli and response sheets. The experimenter paced subjects through this booklet as the subjects performed each task in turn. For each task the following events occurred. Instructions for the task were played from cassette tapes as the subjects read along. Following the instructions, the experimenter answered questions and administered a practice sample of the task. This practice sample served the function of introducing the subject to the task and providing some "warm-up" prior to the critical task. Performance on the practice task was monitored to ensure that subjects understood what the task would require. The critical task was administered upon completion of the practice sample.

Data Analysis

The general analytical plan consisted of two stages. The first stage was primarily concerned with analyses of the individual tasks. Each task was examined to determine whether the expected (from previous findings) or hypothesized (based on "new" treatments) phenomena actually occurred. This first stage was, in essence, a "forms check" for our particular implementations.

As a general analysis, analysis of variance (ANOVA) was used in this stage. The purpose of the ANOVA was to describe and confirm the previous findings on each task, namely, the pattern of significant and nonsignificant effects of the treatments on overall task performance. In addition, since some of the tasks included repetitions within a day, it was possible to test treatment-by-subject interactions. Significant treatmeny-by-subject interactions would mean that subjects responded differently to the treatments; therefore, this interaction effect would indicate that further study would be required in order to identify two or more parameters for use in describing subjects' differences.

An ANOVA for each task was performed both on the raw data and where appropriate, on the transformed scores. The reason for the data transformation was that some of the tasks had a limited range of possible reaponse scores. On those tasks, means and variances would be correlated, which would violate the assumption of ANOVA. Moreover, the limited ranges were likely to produce interaction effects among treatments. For those tasks, arcsin or log transformations were used (Winer, 1962). The reason for analyzing the raw data was to increase understanding of parameters derived from the raw data. Previous research has derived parameters from the raw data and not from the transformed scores. Therefore, it was necessary to perform ANOVAs on the raw data in order to compare the results to those from previous studies.

The second stage of the analysis was to estimate individuals' parameters on each task, such as slopes and intercepts, based upon the results of the ANOVAs. The selection of parameters to be estimated (e.g., slopes and intercepts) was dependent upon significant effects from the ANOVAs. After estimating parameters, they were correlated with each other. In theory, the pattern of correlations would show higher correlation coefficients among those parameters which involve the same information processing operations, thereby providing evidence for the construct validity of the operation.