1-1-1976

Visual search for letters through prose and scrambled prose.

Robert M. Schindler
University of Massachusetts Amherst

VISUAL SEARCH FOR LETTERS THROUGH PROSE AND SCRAMBLED PROSE

A Dissertation Presented

By

ROBERT M. SCHINDLER

Submitted to the Graduate School of the
University of Massachusetts in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 1976

Psychology

VISUAL SEARCH FOR LETTERS THROUGH

PROSE AND SCRAMBLED PROSE

A Dissertation Presented

By

ROBERT M. SCHINDLER

Approved as to style and content by:

Arnold Well, Chairperson of Committee

Alexander Pollatsek, Member

Marvin Daehler, Member

Barry Wadsworth, Member

Bonnie R. Strickland, Chairperson
Psychology Department

ABSTRACT

Visual Search for Letters through

Prose and Scrambled Prose

(August 1976)

Robert M. Schindler, B.A., University of Pennsylvania
M.S., University of Massachusetts
Ph.D., University of Massachusetts

Directed by: Professor Arnold Well

When subjects are instructed to search a prose passage for every occurrence of a target letter, they will tend to make many omission errors. The results of the present study indicate that this tendency is not due to the meaningfulness of syntax of prose, since neither the number nor the distribution of omission errors in scrambled prose passages was found to generally differ from that found in prose passages. However, for both prose and scrambled prose, the higher a word's frequency, the higher was the probability that the subject would miss a target letter occurring in that word. These results suggest that the perceptual processes of reading are elicited by word sequences which are arranged to have the visual features of prose, and that these processes use visual units which are larger than single letters, but not larger than single words.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

C H A P T E R I

INTRODUCTION

Overview

Perhaps one of the main reasons we can function as intelligent organisms is because we can deal with the world as a collection of discrete objects and meaningful patterns. For instance, when we enter a room, we can see at a glance that there are chairs, a table, a lamp, etc., and so we are then prepared to immediately formulate an appropriate course of action.

It is rather unlikely that we are born with the ability to deal with such things as chairs, since chairs have appeared in our environment relatively recently with respect to the speed of genetic adaptation. It is much more likely that we develop such capacities over the course of our experience with the world. Moreover, there are many things which we cannot deal with as single units. If, instead of a chair, we are confronted with a strange and complex piece of machinery, it is probable that a long sequence of inspections and manipulations of its parts would be necessary before we would be able to deal with it in an appropriate fashion. Thus the question of how we come to deal with the world in terms of objects and meaningful patterns is a central one for the understanding of the mind.

One very general explanation for how we come to deal

with our environment in terms of objects and patterns is one which could be called the "automatization hypothesis," and is similar to some of the theoretical notions of Piaget (1969) and Miller, Galanter, and Pribram (1960). When you first deal with a certain novel environmental situation, there occurs a perhaps long series of thoughts and percepts. If these thoughts and percepts are termed "mental acts" (where each "act" is considered a complete moment of consciousness), one might imagine that a similar sequence of mental acts will occur every time you repeat the interaction with that particular environmental situation. The automatization hypothesis holds simply that the common elements of these sequences of mental acts eventually come to occur as a rapid and automatic response to that situation. This automatized sequence becomes an immediate reaction to a situation, which gives that situation meaning, and the presence of the automatized sequence allows the situation to become, phenomenonally, an object or a single pattern.

One prediction of the automatization hypothesis is that if you repeatedly deal with a certain stimulus array for a given purpose, the common elements of the repeated mental sequences will be those features of the array that are most relevant to the task. Then eventually, as repetition continues, these features will become the salient features of how you actually perceive that stimulus, and the features of the stimulus which are irrelevant to the task will corres-

pondingly become difficult to perceive.

A situation where this seems to occur is in reading. When you are first learning to read, you look at every letter separately and slowly sound out each word. But an experienced reader is rarely conscious of most of the individual letters on the page. The commonly observed tendency to fail to notice the misspellings in prose materials (e.g., O'Neill & Ruder, 1974) has even led to the coining of a term, "proofreader's error." Crosland (1924), in perhaps one of the earlist systematic studies of errors made while searching prose, found proofreader's error to be a widely occurring phenomenon. In fact, he found that even experienced proofreaders missed as many as 10% of the misspellings in prose passages which were especially stocked with misspellings.

More recently, Corcoran (1966, 1967) and Corcoran and Weening (1968) have reported observations of errors made while searching prose for the presence or absence of letters. Schindler and Jacobs (1976) found that the number of such errors often decreases markedly when the words of a sentence are rearranged so as to destroy the meaning of the sentence as well as to violate the rules of English syntax, thus providing some evidence that proofreader's error may indeed result from the automatization of mental acts involved in the repeated perception of prose during reading. During these repeated perceptions, the identities of the more redundant words can often be inferred from the context of the passage

so that many of the individual letters of the words are not included in the common elements of the repeated mental acts involved in reading prose. Thus, when the mental acts involved in reading become automatized, these letters become among the least salient features of the stimulus array.

The present experiment was an attempt to use a task similar to proofreading to provide evidence for the automatization hypothesis. Specifically, an attempt was made to replicate and extend the findings of Schindler and Jacobs that subjects will more often fail to detect occurrences of a target letter when it occurs in prose than when it occurs in scrambled prose. And, since the automatization hypothesis would predict that the less important a letter is to the meaning of a sentence, the less salient it will eventually become, the present experiment also investigated whether the prose-specific failures to detect the occurrence of a target letter are correlated with the perceived importance of the word in which the target letter occurs.

Previous Research

In a recent study of visual search for the presence of a letter in prose material, Corcoran (1966) instructed subjects to "go through" a prose passage and cross out all occurrences of the letter e as rapidly as possible. He found (1) that subjects more often failed to cross out the es in words where the e was silent than in words where the e was

pronounced, and (2) the e in the was missed more often than any of the es in other words. Corcoran (1967) found the same pattern of results when he presented subjects with a passage with many letters omitted and instructed them to mark the places where the letters were missing. This similarity of results of a visual search task and a proofreading task suggest that both tasks require many of the same visual processes.

Corcoran failed to find a reliable effect of position of the e in a word, and so concluded that the difference in the detectability of silent and pronounced es occurred because "the acoustic image [of the words] is examined for characteristics normally associated with the presence of an e in the printed word and that the e is more likely to be missed if an acoustic correlate is lacking." However, since the e in the is pronounced, Corcoran suggested that the high probability of missing it in a search task occurred because "the is a highly redundant word, which may be 'taken for granted' and thus not scanned" (Corcoran, 1966).

Krueger (1969, 1970) constructed 12 prose passages so that one of six target letters occurred only once in each passage. Then he created nonprose passages by scrambling the words of each of the 12 prose passages so as to leave the position of the target-containing word unchanged. Six prose and scrambled prose passages which did not contain one of the six target letters were used as "catch" trials.

Krueger's subjects were instructed to search for the target letter as quickly and as accurately as possible. Not only did they search the prose more rapidly than the scrambled prose, but they made fewer errors when searching prose (6.7% errors) than when searching scrambled prose (13.9% errors). When questioned afterwards, Krueger's subjects reported seeing the words as wholes at least 75% of the time, reading 80% of the time, and subvocalizing 55% of the time.

Healy (1976) attempted to generalize Corcoran's finding of an especially high number of visual search errors in the word the by having subjects search a 100-word prose passage for ts rather than es. She instructed her subjects to "read each passage at their normal reading speed" and circle each occurrence of the letter t. She further instructed her subjects not to go back if they realized they missed a t and not to slow down their reading speed "in order to be overcautious about getting the ts." Her results confirmed Corcoran's, since 62% of the omission errors were in the ts of the while only 27.5% of the 40 ts in the passage were contained in the word the.

To test Corcoran's conclusion that the preponderance of erors made in the is due to its redundancy, Healy rearranged the order of the non-t-containing words in the passage to produce a scrambled prose passage with the positions of the 40 occurrences of the target letter unchanged. She reasoned that if subjects miss the ts in the because of the redundancy

of the in prose, the percentage of errors in the should be lower in the scrambled prose passage where one would imagine that all words would have equal importance (or unimportance). Her results showed that the percentage of errors made in thes was just as high in the scrambled prose passage as in the prose passage, and she thus concluded that redundancy is not the crucial variable responsible for the high number of search errors in the.

Healy's results also showed that subjects searched the prose passages faster than the scrambled prose passages, and this result agrees with Krueger's finding. However, unlike Krueger, Healy found subjects made more errors when searching prose (17.3% errors) than when searching scrambled prose (13.3% errors). This discrepancy of results could be due to differences in the instructions given to the subjects of each experiment (search rapidly vs. read at normal speed), or to the fact that Krueger's task was to search for one occurrence of the target letter while Healy's task was to search for many occurrences of the target letter.

To test the possibility that acoustic factors are responsible for the tendency to miss the t in the, Healy conducted another experiment. In this study, search for ts through a scrambled prose passage was compared with search through a passage similar in all ways except that all of the thes were changed to thys. This change left the pronunciation of the t unchanged, but greatly changed the distribution of errors;

the ts in thes constituted 52% of the errors in the passage, while the ts in thy constituted only 8% of the errors made in its passage.

While this experiment ruled out the acoustic hypothesis, it suggested to Healy that word frequency may be a critical factor, since the is the most common word in the language, but thy is a relatively rare one (occurs 12 times/million words, according to Kucera & Francis, 1967). To test if word frequency plays a role in visual search, Healy constructed a 100-word passage containing only nouns, but retaining the pattern of punctuation of the earlier passages. Twenty occurrences of t were in high frequency nouns and twenty were in low frequency nouns. The high and low frequency nouns were matched as to length and position of the target letter in the word. Though this passage was not a scrambled prose passage, it must have looked quite similar to one and the subjects who searched it were given the same "read" instructions as were used in the first experiment. The results showed that ts occurring in high frequency nouns were more likely to be missed than those occurring in low frequency nouns, thus indicating that word frequency plays a role in omission errors made in visual search of word sequences.

Thus Healy concluded that the tendency to miss a target letter occurring in the word the when searching prose is due neither to acoustic factors nor the relative unimportance of

the word the when in prose. Rather, she concluded it is due to the tendency of frequently occurring words to be "read as a unit or chunk rather than in terms of its component letters" (Healy, 1976, p. 235). This will be termed the "unitization hypothesis."

Schindler and Jacobs (1976) presented subjects with a 41- and 46-word paragraph and instructed the subjects to "circle [the target] letter every time it occurs in the lines of print." For each paragraph, the subjects were given 30 seconds to find the ten occurrences of the target letter (f or n). An average of 8.7% of the target letters were missed, but these omission errors were not evenly distributed over the 20 target letter occurrences. A significant correlation ($r = .59$) was found between the number of omission errors made on a target occurrence and the frequency of the target-containing word, thus supporting Healy's conclusion that word frequency affects omission errors made in visual search of word sequences.

However, Schindler and Jacobs also constructed three sentences, each with a similar pattern of six occurrences of three different target letters (f, n, or s). The three sentences, and corresponding scrambled sentences formed by rearranging the non-target-containing words, can be seen in Figure 1. Half of the subjects searched the three sentences and the other half searched the three scrambled sentences, and all subjects were given 30 sec. to search each sentence

Finished files are the result of years of scientific study combined with the experience of many years	Finished files years study the of with of scientific result many are years experinence the of combined
F-sentence	F-nonsentence
Now nearly all of those people in homes in reconstructed areas are happy to be on borrowed time	Now nearly areas people happy in be to in are reconstructed time borrowed those on of homes all
N-sentence	N-nonsentence
Silk screening can be performed and it is vigorously claimed to be as enjoyable as painting for many	Silk screening for and be claimed to it is vigorously many painting as performed as enjoyable be can
S-sentence	S-nonsentence

Figure 1. The sentences and nonsentences used by Schindler and Jacobs (1976). Each of the six target letters in each sentence or nonsentence is circled.

or scrambled sentence. The instructions were the same as those used for the paragraphs.

Overall, 29.5% of the target letters were not circled, and all but two of these omission errors occurred in the short (2-letter) target-containing words. The distribution of these errors can be seen in Table 1. The fs in the ofs were missed more often when they were embedded in a sentence than when they were embedded in a scrambled sentence, and the ns in the ins and on were also missed more often when they were embedded in a sentence. In other words, to a large extent, the errors in the F- and N-sentences were prose-specific. The meaningfulness of the surrounding context did not affect the number of ss which were missed, so these errors were not prose-specific. Since there were prose-specific errors (PSEs) in the N-sentence, the possibility that the prose context increased the number of errors by increasing the extent to which an acoustic representation was scanned appeared unlikely. Instead, noticing that of, in, and on are prepositions and are relatively redundant, while is and as are not prepositions and seem to be more important to the meaning of the sentence, Schindler and Jacobs hypothesized that the importance of the target-containing word was the critical variable behind its errors. If a target letter is in meaningful prose, it will be missed in a visual search task to the extent that it is unimportant to the meaning of the prose.

Further, casual observation and postexperimental ques-

	F	N	S
Sentences	2.04	1.61	.22
Nonsentences	.65	.48	.30

Table 1. Mean number of targets missed per subject in the Schindler and Jacobs study.

tioning suggested to Schindler and Jacobs that PSEs were not caused by voluntary strategy differences; the mechanism which allocates more visual attention to the more important letters seemed to be an automatic one. One might imagine that there are "subroutines" which automatically allocate attention to the letters according to their importance for understanding the prose and that these subroutines are themselves automatically elicited by word sequences conforming to English syntax. For instance, there might be many different subroutines for allocating attention within a prepositional phrase. The particular subroutine elicited would depend on the previous subroutines elicited (i.e., the semantic and syntactic context). Since such automatically elicited subroutines could be considered the primitive beginnings of multiword units, Schindler and Jacobs' "importance hypothesis" is a specific form of the more general automatization hypothesis mentioned in the previous section.

The importance hypothesis implies that not only will unimportant letters get less visual attention than they would in a nonprose context, but also that very important letters may get more visual attention than they would (on the average) in a nonprose context. In other words, a prose context may increase the detectability of some letters. Thus, the importance hypothesis suggests a potential explanation for Krueger's (1970) finding that letter search was both faster and more accurate through prose than through nonprose—

his target letters may have been very important in those particular passages. One rough measure of the importance of a target letter is the frequency of occurrence of the word which contains it. In general, the rarer a word, the more information it contains, and the more important it is to the meaning of the passage. Looking up the Kucera-Francis frequency for each of Krueger's 24 target words shows that the median word frequency is 44 occurrences/million words, and the highest frequency of any word is 1815/million. It is unclear whether these words are rare enough to indicate importance sufficient to cause better detectability in prose, but the absence from Krueger's set of target-containing words of very common words such as the, if, in, and on is worth noting. Also it is worth noting that all of his passages were printed entirely in uppercase letters.

The importance hypothesis also implies that targets in high frequency words should be missed more often only when they are in a prose context; in nonprose frequent words and rare words are equally important (or unimportant). Thus, Healy's finding that the ts in the are missed very often even when in scrambled prose and her finding that frequent nouns are missed more often than infrequent nouns in nonprose both contradict the importance hypothesis. But, on the other hand, Healy's unitization hypothesis cannot account for Schindler and Jacobs' finding that more errors were made in the frequent words of, in, and on when they occurred in

sentences than when they occurred in nonsentences. Such a finding suggests that the largest units for reading may be larger than single words.

Plan of the Experiment

The main purpose of the present study was to confirm and extend the findings of Schindler and Jacobs (1976) concerning the existence of prose-specific errors in visual search through word sequences. There were three specific questions involved:

(1) The first question was whether or not subjects would make more omission errors when target letters were embedded in prose than when they were embedded in scrambled prose. Schindler and Jacobs (1976) and Healy (1976) both found that subjects made more errors when they were searching prose, but Krueger (1970) found the opposite results. Moreover, while Healy's subjects made more errors in the prose passage, they also searched the prose more rapidly, thus raising the possibility that they voluntarily searched the prose passage less carefully (i.e., with less voluntary attention) than they searched the scrambled prose. In the present study each subject searched five prose passages and five scrambled prose passages. An attempt was made to use a wide and representative variety of passages, target letters, and target-containing words. In the hope of equalizing the amount of voluntary

attention given to the prose and scrambled prose passages, subjects were allowed the same amount of time to search the prose and nonprose passages and were instructed to use the entire time period. The importance hypothesis predicts that more errors will be made in prose, but only if it is assumed that there are generally more letters which are made less important by a prose context than are made more important. While this assumption seems reasonable (since there is so much redundancy in prose), its presence does mean that the comparison of the total number of errors made in prose and scrambled prose is not a decisive test of the importance hypothesis.

(2) The second question was whether the distribution of errors would differ with the meaningfulness of the context. The importance hypothesis predicts that since the errors in only the prose-embedded targets will vary with importance, the distribution of errors in the prose passages should be different from that in the scrambled prose passages. Assuming the subjects read both the prose and scrambled prose passages, the unitization hypothesis predicts that the distribution of errors should not differ between prose and scrambled prose. Even if, overall, more errors are made in prose passages, a word which is a single unit (or two units, etc.) should continue to be so regardless of the meaningfulness of the passages, and should not change because of the

interword dependencies of a multiword unit. Thus each word should cause an unchanging proportion of the total errors made in the passage.

(3) The third question is, assuming that the distribution of errors differs between prose and scrambled prose passages, whether these PSEs (errors for a target occurrence when in prose minus errors for the same target occurrence when in scrambled prose) will correlate with the perceived importance of the target-containing words in prose. The importance hypothesis predicts that PSEs and importance will be negatively correlated. In the present study the perceived importance of the target-containing words was obtained by having the subjects who searched the passages also rate the importance of all the words in each of the five prose passages they received. Half of the subjects rated importance by crossing out the 20 least important words in the passages (Telegraph method) and half of them rated importance by direct estimation using a 1-5 scale (Estimation method).

In addition to treating these three specific questions, the present study was designed to gauge the extent to which any effects of passage meaningfulness were under voluntary control. To do this, half of the subjects were instructed to read their prose and scrambled prose passages and the other half were instructed to not read, but rather to search the words letter by letter. If passage meaningfulness effects

found using the Read instructions do not occur with the Search instructions, then such effects would result from differences in choice of strategy for searching prose and scrambled prose passages. But if meaningfulness effects remained even under Search instructions, it would suggest that the processes used for searching meaningful word sequences are automatic (and thus, in at least that sense, perceptual).

In order to assess the extent to which subjects read the prose and scrambled prose passages, they were given a questionnaire, and were also given a short word recognition test immediately after searching the passages. The recognition test included two non-target-containing words from each of the passages searched, so, assuming that reading the words of a passage is necessary for correct recognition of the words (cf., Craik & Lockhart, 1972), the score on the recognition test was a second measure of the extent to which subjects read the passages.

Although both Healy (1976) and Schindler and Jacobs (1976) provide evidence against the hypothesis that acoustic representations are being searched in this task, the subjects were asked how often they utilized the acoustic properties of the target letter while searching the passages. Also, miscellaneous "biographic" information, such as estimated reading speed, age, handedness, etc. was obtained, so that it would be possible to test for any correlations of these variables with omission errors made in the search task.

CHAPTER II

METHOD

Materials

Each subject received an 18-page booklet which contained (1) 10 passages[1] to be searched, (2) 5 prose passages to be rated for importance, (3) a recognition test, and (4) a questionnaire.

The ten prose passages were constructed as follows: First, the letters A, E, I, O, F, H, N, S, T, and W were chosen to be target letters. An attempt was made to choose those letters which occur in the most frequent words (e.g., the, of, in, was) and to have a mixture of both vowels and consonants, but otherwise the selection was arbitrary. Then, issues of Reader's Digest (1970-71) were scanned for passages which contained many occurrences of one of the target letters but few proper nouns, symbols, or any other forms which could complicate scrambling or add unnecessary variation to the words in the passages. Then the passages were altered so that they contained exactly 60 words and at least 20 occurrences of one of the ten target letters. An effort was made

---

[1]The term passage will be used in this paper to designate any of the 60-word arrays used in the study. Half of the passages used were prose and half were scrambled prose. The term passage type will be used to designate a specific set of 60 words. For instance, the prose passage where A was the target letter and the scrambled prose passage where A was the target letter are both examples of the A-passage type.

to avoid having target letters occur in consecutive words.

The ten scrambled prose passages were constructed by rearranging the non-target-containing words of each prose passage to produce as meaningless and as syntactically illegal a word sequence as possible. Care was taken to not change the position of each target-containing word on the line and on the page. For example, if two target-containing words in a prose passage were separated by a 5-letter word, they would be separated by a 5-letter word, two 2-letter words, or a 3-letter word and a 1-letter word in the corresponding scrambled prose passage.[2] This is of course why consecutive target-containing words were avoided; they could not be scrambled relative to each other without changing their position in the passage.

Each of the 10 prose passages and 10 scrambled prose passages was typed using an Olympia electric typewriter with pica type, and then was mimeographed onto white 8-1/2" by 11" sheets of paper. Each passage was double-spaced and centered on the page. The upper- and lowercase form of the target letter for the passage was printed at the top of the page. The first line of both the prose and scrambled prose passages was indented five spaces, and the pattern of capitalization and punctuation of the prose passages was retained in the

---

[2]Occasionally, the spacing between two target-containing words had to be altered by a space or two in order to get an adequately scrambled passage.

scrambled prose passages. The overall result was that each scrambled prose passage differed from its corresponding prose passage only in the arrangement of the non-target-containing words. A copy of each of the 20 passages in the form in which they appeared in the study can be seen in Appendix A.

For each prose passage two rating sheets were constructed, one for the Telegraph rating method and the other for the Estimation method. A Telegraph rating sheet consisted of the prose passage (typed in the same form as on the sheet to be visually searched) printed once on the upper half of the page and once on the lower half of the page. An Estimation rating sheet consisted of the prose passage printed in double-spaced paragraph form at the top of the page. But at the bottom of the page, the passage was printed with extra spaces between the words and the lines, and under each word was a short line on which the subject could write his numerical rating of the importance of each word. The rating sheets were typed and mimeographed in the same way as were the sheets which were searched. See Appendix B for examples of a Telegraph rating sheet and an Estimation rating sheet.

The recognition test was constructed by selecting two relatively distinctive non-target-containing words from each of the ten passage types. Each of these 20 words was matched with two words of equal length and comparable frequency of occurrence in printed English. The word taken from a passage

and its two "distractor" words were typed in a row, so that the recognition test consisted of 20 such rows. There were two versions of the recognition test, each with a different pseudorandom order of the twenty rows, and each with a brief instructional paragraph at the top of the page. The recognition test sheets were typed and mimeographed in the same way as were the other sheets. A copy of the recognition tests can be seen in Appendix C.

The questionnaire consisted of two separate sheets. The first sheet, also serving as the cover sheet for the booklet, asked the subject's age, sex, dominant hand, year in school, and asked the subject to estimate how much he reads and how fast he reads. The second sheet was placed after the passages to be searched and the recognition test, but before the rating sheets. It asked the subjects to estimate (1) how often they had read the paragraphs, (2) how often they had read the scrambled paragraphs, (3) how often they had used the acoustic properties of the target letter, and (4) whether they had found it more difficult to find the targets in the prose or the scrambled prose passages. Copies of the two sheets of the questionnaire can be seen in Appendix D.

Design

The cover page for all of the booklets was the first questionnaire sheet. Pages 2 through 11 were the ten passages to be searched. Each of the ten passage types were

represented, but a pseudorandom procedure determined whether any given passage occurred in its prose or scrambled prose form. Page 12 was the recognition test. Page 13 was the second questionnaire sheet, and pages 14 through 18 were rating sheets for the five prose passages which were searched.

Since instructions (Read vs. Search) and method of rating word importance (Telegraph vs. Estimation) were between-subject variables, the subjects were equally divided among the following four conditions: Read-Telegraph, Read-Estimation, Search-Telegraph, and Search-Estimation. For the booklets used in the Read-Telegraph condition, pages 14 through 18 were Telegraph rating sheets in pseudorandom order. For all of the booklets in the Read-Telegraph condition the passages to be searched occurred in alternating prose-scrambled prose order, but in half of the booklets a prose passage began the sequence and in the other half a scrambled prose passage began the sequence. Further, among the booklets in the Read-Telegraph condition, each passage type occurred half the time in its prose form and half the time in its scrambled prose form. For the booklets used in the Read-Estimation condition, pages 14 through 18 were Estimation rating sheets in pseudorandom order. The passages to be searched in the Read-Estimation booklets were constrained in the same ways as were those in the Read-Telegraph condition.

Each booklet in the Search-Telegraph condition was matched with a booklet in the Read-Telegraph condition, and

each booklet in the Search-Estimation condition was matched with a booklet in the Read-Estimation condition. This was done in order to minimize any differences between the Read and Search conditions other than the differing instructions.

Procedure

Each subject began by answering the questions on the cover sheet of the booklet. Then instructions were read (see Appendix E for all instructions read during the experimental session). Read subjects were told to "read the sequence of words and circle the target letter every time it occurs," and Search subjects were told to "search each word letter by letter, and totally ignore its meaning." Subjects were given one minute to search each page, since it was determined from pilot subjects that this would be more than enough time for most people. Subjects were encouraged to check their work if they finished before the minute was up. The interval between pages was 10-15 seconds.

Immediately after completing the tenth passage, the subjects were instructed to turn to the recognition test and to circle the one word in each row which had occurred in the preceding passages. Subjects were given as long as they wanted to complete the recognition test. Following the test, they completed the second questionnaire sheet.

Next, instructions for the Telegraph or Estimation rating methods were read. Subjects were given five minutes to

complete the ratings on each passage. However, since it turned out that many subjects required much less than five minutes, the instructions were changed to allow subjects to go from passage to passage at their own pace.

Subjects

Two hundred and two University of Massachusetts undergraduates, one graduate student, and one research assistant served as subjects. Four of the initial 200 subjects had to be discarded. One was discarded because English was not her native language, one was discarded because he failed to follow instructions, and two were discarded because they obviously failed to complete the searching of one or more of the passages. Replacements for the discarded subjects were run so that there were 50 subjects in each of the four between-subject conditions. Subjects were run in groups for one 45-minute session, and received course credit for their participation.

CHAPTER III

RESULTS

Overall Results

The 200 subjects who each searched ten passages made a total of 3,569 errors. No subject failed to make at least one error. Almost all of these errors (3554) were omission errors (failure to circle the occurrence of a target letter), and so from here on, omission errors will often be termed simply "errors." However, there were also 15 commission errors (circling a letter other than the target letter), and these will be discussed separately in a later section.

Since 3554 omission errors were made by 200 subjects, the average subject made 17.77 errors. Since there was a total of 260 occurrences of target letters in the 10 passages, the average subject's error rate was 6.83%. This figure is rather low compared to the error rates found in previous studies of visual search through prose and scrambled prose. Krueger's (1970) subjects searched prose and scrambled prose displays for one occurrence of a target letter and showed a mean error rate of 10.0%. Healy's (1976) subjects searched prose and scrambled prose for 40 occurrences of a target letter and missed an average of 15.25% of them. However, both Krueger's and Healy's subjects were under a certain amount of time pressure. The lower error rate found in the present study may be due to the fact that most subjects were given

more time than they would have taken if under time pressure, and many did in fact have time to go back and check their work.

The mean error rates on the prose and scrambled prose passages for the Read and Search subjects can be seen in Table 2.[3] In order to correct for deviations from normality of data in percentage form, the arcsin transformation of the error rates for each subject was computed. An analysis of variance on these error rates showed there to be no significant effect of instructions ($F(1,198) < 1$), no significant difference between the error rates in prose and scrambled prose passages ($F(1,198) = 3.00$, $p > .10$), and no significant interaction between instructions and the meaningfulness of the passage ($F(1,198) = 1.02$, $p > .20$). However, although none of these effects are statistically significant, the trends in the data are reasonable and consistent: there was a tendency for the error rates to be higher in prose passages than in scrambled prose passages, and this difference tended to be larger for the subjects who received Read instructions. Since the trends are reasonable, the lack of statistical significance could be due to inconsistencies between subjects as to the size and direction of effects. With this in mind, the

---

[3]Analyses were performed on the error rates rather than on the number of errors since the particular passage types which were in prose or scrambled prose form differed between subjects, and thus the number of possible errors for the prose and scrambled prose conditions differed between subjects.

Table 2

Mean Error Rates on Prose and Scrambled Prose Passages for Subjects given Read and Search Instructions

	Read Instructions	Search Instructions
Prose	7.63	6.93
Scrambled Prose	6.41	6.67

analysis of variance was repeated using passage type rather than subjects as the random effects variable, but the effect of instruction, of meaningfulness, and their interaction still failed to reach accepted levels of statistical significance ($F(1,9) < 1$; $F(1,9) = 2.82$, $p < .20$; $F(1,9) = 3.77$, $p < .10$).

Since the recognition test contained two items from each passage, it was possible to give each passage a recognition score. The mean recognition scores for prose and scrambled prose passages for Read and Search subjects can be seen in Table 3. An analysis of variance using subjects as the random effects variable indicated that more words from the prose passages were recognized than from the scrambled prose passages ($F(1,198) = 20.86$, $p < .001$), but that there was no significant effect of instructions on recognition scores ($F(1,198) < 1$), and no significant interaction between the effect of passage meaningfulness and instructions ($F(1,198) = 2.07$, $p < .20$). An analysis of variance using passage type as the random effects variable also indicated that prose passages had higher recognition scores ($F(1,9) = 6.38$, $p < .05$), but indicated that the higher mean recognition score for the Read subjects than for the Search subjects was reliable over the 10 passage types ($F(1,9) = 10.62$, $p = .01$). Although the effect of passage meaningfulness was larger for the Read subjects (as was the case for errors rates) and the effect of instructions was apparent for only the prose passages, the

Table 3

Mean Number of Correct Recognitions per Passage (out of two possible) for Prose and Scrambled Prose Passages and for Read and Search Instructions

	Read Instructions	Search Instructions
Prose	1.34	1.26
Scrambled Prose	1.16	1.16

analysis of variance indicated that this interaction was not reliable over passage types ($F(1,9) = 1.97$, $p < .20$). In sum, the effect of passage meaningfulness on recognition score was highly reliable, but, apparently, the Read instructions increased the recognition score for only some of the subjects.

Although it was not found that the meaningfulness of the passages affected the overall number of omission errors made, it is possible that meaningfulness did cause a change in how the omission errors were distributed among the 20-32 occurrences of the target letter in a passage. The distribution of omission errors for the prose and scrambled prose forms of each of the ten passage types can be seen in Table 4.

A two-way analysis of variance was performed on the errors for each of the ten passage types. This was felt to be an acceptable procedure even though the data points were dichotomous (a subject either did or did not make an error in a given target letter position) and the distributions were extremely skewed (most of the subjects did not make an error at a given position), since there is evidence that even such flagrant violations of the assumption of normality cause only small distortion of alpha (Glass, Peckham, & Sanders, 1972, pp. 250-251). However, since the variable of target letter position contained as many as 32 levels, it was felt that a correction for possible non-homogeneity of covariance was necessary. Thus, the F's obtained were evaluated

Table 4

The Number of Omission Errors at Each Target Position for the Prose and Scrambled Prose Forms of Each of the Ten Passage Types

Target Letter		Target Letter Position													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	Prose	0	0	4	1	5	9	2	0	2	2	5	1	7	7
	Scrambled Prose	0	0	7	1	8	12	3	2	1	1	6	1	3	1
E	Prose	0	0	0	17	4	5	2	9	3	5	1	4	16	17
	Scrambled Prose	0	1	3	10	3	9	1	5	0	7	0	9	18	22
I	Prose	0	0	0	0	0	14	0	0	1	2	0	0	3	22
	Scrambled Prose	0	2	1	1	1	5	1	0	1	0	0	0	4	26
O	Prose	2	0	3	7	0	7	1	25	0	0	2	0	10	1
	Scrambled Prose	0	0	2	1	0	7	2	33	0	1	3	0	1	1
F	Prose	1	76	0	1	3	3	37	1	0	7	61	0	6	4
	Scrambled Prose	0	55	0	0	1	2	55	1	0	2	67	0	8	2
H	Prose	13	24	15	11	3	15	2	42	9	1	2	36	17	39
	Scrambled Prose	7	25	19	14	3	16	5	29	13	3	2	35	22	29
N	Prose	0	0	1	11	4	0	3	5	16	6	1	38	18	2
	Scrambled Prose	1	1	0	13	2	2	0	2	13	11	2	50	9	5
S	Prose	1	9	2	2	1	19	1	1	0	0	14	0	4	1
	Scrambled Prose	1	5	0	0	0	3	2	2	1	1	5	1	4	1
T	Prose	8	1	0	6	3	9	1	21	2	0	0	1	0	0
	Scrambled Prose	2	1	0	4	2	9	1	16	1	0	0	2	0	0
W	Prose	0	2	1	3	0	9	1	25	10	2	0	18	1	6
	Scrambled Prose	0	6	1	1	3	4	1	9	11	2	0	4	2	5

Table 4 (continued)

Target Letter		Target Letter Position														
		15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
A	Prose	3	6	9	2	23	2	2	17	12	12	15	1			
	Scrambled Prose	7	4	6	1	34	3	2	6	15	1	8	1			
E	Prose	4	16	7	1	1	1	36	0	5	1	0	4	3	1	4
	Scrambled Prose	13	5	4	3	2	0	27	0	4	0	1	1	2	5	3
I	Prose	0	2	2	11	2	2	1	0	4	0	0	0	1		
	Scrambled Prose	1	5	0	22	1	8	2	0	4	2	0	0	0		
O	Prose	2	0	0	1	1	4	2	0							
	Scrambled Prose	3	0	0	0	0	1	6	0							
F	Prose	1	1	19	0	0	12									
	Scrambled Prose	0	0	19	2	0	18									
H	Prose	10	14	11	2	10	6	6	50	33	29					
	Scrambled Prose	8	21	25	4	12	7	20	29	34	27					
N	Prose	17	0	0	18	2	5	6	2	0	35	6	23	1	33	11
	Scrambled Prose	27	0	2	21	0	4	6	7	1	27	1	42	4	30	6
S	Prose	6	2	2	2	14	2	2	1	1	5	34	1	0	4	5
	Scrambled Prose	10	5	5	1	20	6	5	1	1	6	9	1	0	5	4
T	Prose	14	0	0	39	12	1	2	1	1	2	25	3			
	Scrambled Prose	12	0	1	8	3	1	1	1	1	1	20	7			
W	Prose	4	3	18	24	0	8	12								
	Scrambled Prose	7	0	9	13	2	7	1								

Table 4 (continued)

Target Letter		Target Letter Position 30	31	32	Totals
A	Prose				149
	Scrambled Prose				134
E	Prose	5	3		175
	Scrambled Prose	3	4		165
I	Prose				67
	Scrambled Prose				87
O	Prose				68
	Scrambled Prose				61
F	Prose				233
	Scrambled Prose				232
H	Prose				400
	Scrambled Prose				409
N	Prose	48	6	5	323
	Scrambled Prose	18	10	21	328
S	Prose	0	0		136
	Scrambled Prose	1	0		106
T	Prose				152
	Scrambled Prose				94
W	Prose				147
	Scrambled Prose				88
					3554

against lowered degrees of freedom, which were calculated from the covariance matrix for each passage according to the correction of Greenhouse and Geisser (Winer, 1973, p. 523). The results of these ten ANOVAs can be seen in the first 5 columns of Table 5. The effect of target letter position is highly significant for all passage types, indicating clearly that each occurrence of the target letter was not equally likely to be missed in the present task. The effect of passage meaningfulness (prose vs. scrambled prose) was significant only for the T and W passages; in these two passage types, more errors were made in the prose forms of the passage than in the nonprose forms. The interaction of passage meaningfulness and target position was significant at the .01 level for all passage types except the I-passage ($p < .05$) and the O-passage ($p < .10$). This indicates that for at least eight out of the 10 passage types, the distribution of the errors among the target positions was affected by whether or not the non-target-containing words formed a prose or a nonprose context.

However, since Healy (1976) found that subjects take longer to search scrambled prose than to search prose, it is possible that the effect of meaningfulness on the distribution of errors found for eight of the passage types could be due to a need for the subjects to rush themselves at the end of the scrambled prose passages more than they needed to rush themselves at the end of the prose passages. Or, if some

Table 5

Results of ANOVAs on the Omission Errors Made for Each Passage Type (M = meaningfulness, TP = target position)

Passage Type	Source of Variation	All TP df_1, df_2	All TP F	All TP p	First 10 TP df_1, df_2	First 10 TP F	First 10 TP p
A	M	1,198	<1	--	1,198	1.21	--
	TP	12,2380	16.28	<.001	4,832	8.75	<.001
	M x TP	12,2380	2.35	<.01	4,832	<1	--
E	M	1,198	<1	--	1,198	<1	--
	TP	12,2459	20.41	<.001	5,1015	9.50	<.001
	M x TP	12,2459	1.78	<.01	5,1015	1.53	<.20
I	M	1,198	1.97	<.20	1,198	<1	--
	TP	6,1261	25.24	<.001	2,468	12.30	<.001
	M x TP	6,1261	2.17	<.05	2,468	3.80	<.02
O	M	1,198	<1	--	1,198	<1	--
	TP	5,1039	20.93	<.001	2,518	44.13	<.001
	M x TP	5,1039	2.09	<.10	2,518	1.66	<.20
F	M	1,198	<1	--	1,198	<1	--
	TP	5,1181	160.77	<.001	2,547	205.29	<.001
	M x TP	5,1181	4.19	<.001	2,547	8.11	<.001
H	M	1,198	<1	--	1,198	<1	--
	TP	14,2805	26.34	<.001	6,1300	23.84	<.001
	M x TP	14,2805	2.57	<.005	6,1300	1.58	<.20
N	M	1,198	<1	--	1,198	<1	--
	TP	13,2700	39.08	<.001	4,857	13.30	<.001
	M x TP	13,2700	5.10	<.001	4,857	<1	--
S	M	1,198	1.77	<.20	1,198	8.17	<.005
	TP	10,2138	15.33	<.001	3,709	11.04	<.001
	M x TP	10,2138	5.38	<.001	3,709	5.87	<.001
T	M	1,198	9.24	<.01	1,198	2.14	<.20
	TP	8,1727	25.40	<.001	3,787	17.52	<.001
	M x TP	8,1727	5.35	<.001	3,787	<1	--
W	M	1,198	6.56	<.02	1,198	2.05	<.20
	TP	9,1952	13.68	<.001	4,898	15.03	<.001
	M x TP	9,1952	3.79	<.001	4,898	4.13	<.005

subjects needed more time to search prose passages for a letter, they may have been more rushed at the end of the prose passages than at the end of the scrambled prose passages. Either phenomenon could cause an interaction between passage meaningfulness and target position which would be an artifact of the present experimental procedure. To test for this possibility, the sum of the absolute values of the prose-specific errors (more errors in either the prose or scrambled context of a target position) for the first half of the target positions in each passage was compared with the sum of the absolute value of the PSEs for the later-occurring half of the target positions in the passage. There were more (positive or negative) PSEs in the later positions in 7 out of the 10 passage types, and the mean number of PSEs in the earlier target positions was 33.7 errors while the corresponding number for the later positions was 51.2 errors. Thus, it is quite possible that differential rates of search for prose and scrambled prose for the later target positions of the passage could be responsible for the interactions found between meaningfulness and target position.

In an attempt to circumvent the possibility of this explanation, the ANOVAs on the errors for each passage type were rerun (again correcting the degrees of freedom for heterogeneity of covariance), this time using data from only the first 10 target positions. These results can be seen on the right-most three columns of Table 4. Although the main ef-

fect of target position remained highly significant, the interaction between meaningfulness and target position remained significant for only three passage types. Thus the present results cannot be taken to support the conclusion that, in general, the meaningfulness of the passage affects the distribution of omission errors in visual search for letters.

To determine if the meaningfulness of the passage affected how the errors were distributed in single words, the 40 target-containing words which contained at least two occurrences of the target letter[4] were examined apart from the other target-containing words. Since each of the 40 words appeared in a prose context as often as it appeared in a scrambled prose context, it was possible to observe whether the meaningfulness of a word's context affected the pattern of visual search omission errors within that word. In general, the second occurrence of the target letter in a word was more likely to be missed than the first occurrence; almost 3/4ths of all the errors made in the 40 words were omissions of the second occurrence of the target letter. If this percentage for the prose-embedded words differed from that for the scrambled-prose-embedded words, then it could be said that the meaningfulness of the context changed the distribution of errors within a word.

---

[4]Two of the 40 words contained three occurrences of the target letter, but the third occurrence was not included in this analysis.

The results were that when the 40 words were embedded in prose, 78.1% of all errors made in those words were made in the second target occurrence. When the 40 words were embedded in scrambled prose, only 66.6% of all errors made in those words were made in the second occurrence of the target letter. This difference was reliable across subjects, $t(199) = 3.97$, $p < .001$.

However, since fewer errors were made on the target-containing words in scrambled-prose contexts (though not significantly fewer), it is possible that many subjects made no errors at all on either of the two occurrences of targets in scrambled prose embedded words, and thus causing the mean percent of the errors that were made in the second position to be artificially low. To correct for such a "floor effect," all subjects who made no errors in either the prose- or scrambled-prosed-embedded target words were excluded from the analysis. However, the remaining 137 subjects still made a greater percentage of their errors on the second occurrence of the target letter ($t(136) = 4.00$, $p < .001$).

Unfortunately, it appeared as though this tendency for subjects to make more of their errors in the second occurrence of the target letter when the target-containing word was embedded in a prose context was not reliable over words, since when words were used as the random effects variable, the prose/scrambled prose difference did not quite reach the accepted level of significance ($t(39) = 1.70$, $p = .08$).

Thus, the present results indicate that the meaningfulness of a word's context can change the pattern of distribution of errors within a word, but that whether such changes occur may depend on the particular word involved and/or on the specific form of the word's context.

Subject Variables

The distribution of the responses of the 200 subjects to the ten questionnaire items can be seen in Table 6. In addition, for all subjects the following two numbers were computed: (1) the total number of correct responses on the recognition test, and (2) the number of correct responses on the recognition test items taken from prose passages minus the number of correct recognition test items taken from scrambled prose passages (termed "prose-specific recognition score", PSR). These 12 quantities (plus the instruction condition of the subject) were all correlated with each other, and the resulting matrix of correlation coefificients can be seen in Table 7. These 13 "subject variables" were also correlated with the total errors and PSEs for each subject,[5] and the coefficients are displayed in Table 8. Inspection of Tables 7 and 8 suggest that the factor most strongly affecting the number of errors a subject made on the ten passages was his

[5]The PSE score for a subject was computed by subtracting the percentage of errors he made when searching his five prose passages from the percentage of errors he made when searching his five scrambled prose passages.

Table 6

Breakdown of the Responses of the 200 Subjects to the 10 Questionnaire Items (see Appendix D for actual wording of the questionnaire)

1. Age:	18 yrs.	19 yrs.	20 yrs.	21 yrs.	22-33 yrs.	
	51	58	29	27	35	
2. Sex:	Male	Female				
	86	114				
3. Dominant hand:	Left	Right				
	25	175				
4. Yr. in School:	Fresh.	Soph.	Jr.	Sr.	Grad.	Other
	83	43	52	20	1	1
5. No/pages read/wk.:	0-25	25-50	50-100	100-200	200+	
	34	44	50	50	22	
6. Reading Speed:	much slower	below average	average	above average	much faster	
	9	32	116	39	4	
7. Read prose passages:	<5%	5-25%	25-75%	75-95%	>95%	
	45	55	55	34	11	
8. Read scrambled prose passages:	<5%	5-25%	25-75%	75-95%	>95%	
	93	66	29	9	3	
9. Used sound of target letter:	<5%	5-25%	25-75%	75-95%	>95%	
	69	44	51	25	11	
10. Difficulty:	prose more diff.		no diff.		prose less diff.	
	57		90		53	

Table 7

Correlation Matrix for 13 Subject Variables. A Correlation Coefficient of .14 Is Necessary for $p < .05$, and .18 for $p < .01$, Two-tailed (df = 198)

Variable[1]	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1	.72	-.21	.07	.04	-.14	.04	-.13	-.16	.05	-.22	.03	.02
2	.72	1	-.20	.08	.10	-.10	.04	-.04	-.04	.06	-.09	.02	.02
3	-.21	-.20	1	-.02	.02	.03	-.07	-.08	-.09	-.01	.07	-.08	-.02
4	.07	.08	-.02	1	.08	.04	0	-.08	.05	-.03	-.03	.05	.01
5	.04	.10	.02	.08	1	-.06	-.10	-.07	-.08	-.07	-.03	.05	.01
6	-.14	-.10	.03	.04	-.06	1	-.01	.25	.16	0	.21	0	-.02
7	.04	.04	-.07	0	-.10	-.01	1	.15	.07	.13	.03	.08	.01
8	-.13	-.04	-.08	-.08	-.07	.25	.15	1	.58	.18	.20	.13	.02
9	-.16	-.04	-.09	.05	-.08	.16	.07	.58	1	.17	.19	.07	.06
10	.05	.06	-.01	-.03	-.07	0	.13	.18	.17	1	.33	.07	-.06
11	-.22	-.09	.07	-.03	-.03	.21	.03	.20	.19	.33	1	.04	.02
12	.03	.02	-.08	.05	.05	0	.08	.13	.07	.07	.04	1	-.03
13	.02	.02	-.02	.01	-.26	-.02	.01	.02	.06	-.06	.02	-.03	1

[1] 1 = Age; 2 = Year; 3 = Sex, 1 = male, 2 = female; 4 = Dominant Hand, 1 = left hand dominant, 2 = right hand dominant; 5 = Instructions; 6 = Total Recog. Score; 7 = PSR: 8 = Read Prose; 9 = Read Scrambled Prose; 10 = Amount Read; 11 = Reading Speed; 12 = Sound Target Letter; 13 = Scrambled Prose Difficulty

Table 8

Correlation Coefficients of Errors with Subject Variables

A Coefficient of .14 Is Necessary for $p < .05$,

and .18 for $p < .01$, Two-tailed (df = 198)

Variable	Total Errors	Prose-specific Errors
Age	.34	-.07
Year	.29	-.06
Sex	-.14	.12
Dom. hand	.05	-.08
Instructions	-.01	-.07
PSR	-.02	.12
Total recog score	-.16	-.05
Read prose	-.09	-.02
Read scrambled prose	-.06	0
Amt. read	.03	.02
Reading Speed	-.18	.09
Sound target letter	0	-.06
Scrambled prose diff.	.16	-.04

age or year in school. A stepwise multiple regression was performed, using total errors as the dependent variable and the 13 subject variables as the independent variables. It indicated that after the variable of age was entered into the regression equation, only the variable of scrambled prose difficulty removed any significant further variation. It is unclear why the variables age/year and scrambled prose difficulty should affect the number of errors made in this task. However, it should be noted that none of the correlations were particularly strong; with all of the 13 variables entered, the regression equation could account for only 18.5% of the total error variance.

Prose-specific errors did not correlate significantly with any of the 13 subject variables. However, among the larger of these nonsignificant correlations was the positive correlation between PSEs and reading speed; it appears there was a tendency for faster readers to make more PSEs. This is especially interesting in light of the fact that reading speed is negatively correlated with total errors ($p = .01$). In order words, the tendency of the fast readers to make more PSEs was not due simply to their making more errors in general.

Since it seemed reasonable that faster readers would have more automatized visual processes for dealing with prose, the relationship between reported estimates of reading speed and PSEs was selected for further examination. To do

this, the 43 subjects who rated themselves as reading at an "above average" speed or "much faster" than average were designated "fast readers." Unlike the total sample of 200, the fast readers made significantly more errors on prose passages than on the scrambled prose passages ($t(42) = 2.66$, $p = .01$). An analysis of variance performed on the arcsin transformed error rates indicated that the 43 fast readers made more PSEs than the 157 average or slow readers ($F(1, 198) = 4.30$, $p < .05$). The means for each group (see Table 9) indicate that the fast readers tended to make fewer errors than average and slow readers on the scrambled prose passages (though not significantly fewer, $F(1,198) = 1.27$, $p > .20$), but about the same number of errors on the prose passages. Taken together with the negative correlation between reading speed and total errors, this observation suggests that fast readers are more accurate than the average in searching for letters through word sequences in general, but that this increase in accuracy does not extend to searching for letters through word sequences which are meaningful.

The mean recognition scores for the fast readers and the other subjects can be seen in Table 10. An analysis of variance performed on these data indicated that the effect of passage meaningfulness was significant ($F(1,198) = 8.36$, $p < .005$), but the effect of reported reading speed and the interaction were not significant ($F(1,198) < 1$; $F(1,198) = 2.32$, $p < .20$). Although the interaction of the effects of

Table 9

Mean Error Rates on Prose and Scrambled Prose Passages for Fast Readers and for Average and Slow Readers

	Fast Readers	Average and Slow Readers
Prose	7.53	7.04
Scrambled prose	5.55	6.81

Table 10

Mean Number of Correct Recognitions per Passage (out of two possible) for Prose and Scrambled Prose Passages for Fast Readers and for Average and Slow Readers

	Fast Readers	Average and Slow Readers
Prose	1.56	1.42
Scrambled Prose	1.25	1.32

passage meaningfulness and reading speed on recognition score was not significant, the trend suggests that the fast readers may have been more likely to read prose passages than were the average and slow readers, but less likely to read scrambled prose passages. To further investigate this tendency, Table 11 presents the questionnaire responses for the read-prose and read-scrambled-prose questions for the fast and average/slow readers. While it appears as though fast readers were more likely to read than average/slow readers and all readers were more likely to read prose passages than nonprose passages, it is unclear whether the questionnaire responses support the trend toward interaction shown by the mean recognition scores.

Word Variables

Each of the 260 target letter occurrences was scaled along the following dimensions:

(1) line on page--the ordinal number of the line on the page in which the target-containing word appeared (varied from 1-8)

(2) column on page—The print on each page was divided into four equal-sized columns, and numbered from left to right. The value of this variable was the number of the column containing the target occurrence.

(3) position in word—the ordinal position of the target letter occurrence in the word (varied from 1-11)

Table 11

The Percentage of Fast Readers and Average/Slow Readers Who Gave Each Response to the Read-Prose and Read-Scrambled-Prose Questions

		% of time spent reading				
		<5%	5-25%	25-75%	75-95%	>95%
Fast Readers	Prose	14	21	35	21	9
	Scrambled Prose	33	35	23	9	0
Average and Slow Readers	Prose	25	29	25	16	4
	Scrambled Prose	50	32	12	3	2

(4) position in sentence—the ordinal number of the target-containing word in the sentence (or clause separated by a comma) divided by the total number of words in the sentence or clause (varied from .04-1.00)

(5) position in passage—the ordinal number of the target letter occurrence in the passage divided by the total number of target letter occurrences in that passage (varied from .03-1.00)

(6) pronounceability—The phoneme representing how the occurrence of the target letter was pronounced in its word context was assessed, based on the word pronunciations given in the American College Dictionary (Bernhart, 1966). The measure of pronounceability used was the number of times that that phoneme corresponded with the graphemic occurrence of the target letter divided by the total number of occurrences of that phoneme in the pronunciations of a 100,000-word sample of printed English (Dewey, 1970). For example, if the target letter "T" was pronounced /t/, it would have a pronounceability rating of .975. This means that if there is a /t/ in the acoustic representation of a word, there is a 97.5% chance that there is a t in the visual representation of the word. If the subjects scanned an acoustic representation of the words and checked the visual features of the word for the presence of the target letter only if they encountered a phoneme that usually corresponded to

the occurrence of the target letter, then target occurrences ranking low on this measure of pronounceability should be often missed.

(7) target in word—the total number of target letter occurrences within the word (varied from 1-3)

(8) word length—the number of letters in the word containing the occurrence of the target letter (varied from 1-14)

(9) telegraph rating—the total number of times the target-containing word was crossed out (for being unimportant) in the Telegraph rating task (to make this scale increase as importance increased, its values were made negative, so it varied from -48 to 0).

(10) estimation rating--the mean Estimation rating for the target-containing word (varied from 2.18-3.96)

(11) letter frequency—the number of occurrences (in thousands) of the target-containing word in a 100,000-word sample of printed English (Dewey, 1970; varied from 9-55)

(12) word frequency—the number of occurrences of the target-containing word in a 1-million-word sample of printed English (Kucera & Francis, 1967; varied from 0-69,971).

Since past research (e.g., Howes & Solomon, 1951) has found word frequency effects to be a negatively accelerated function of word frequency, the log of the word frequency of each target-containing word was computed, to make a total of 13 "word variables." These word variables were correlated with

each other, and the resulting matrix of correlation coefficients can be seen in Table 12.

Inspection of Table 12 shows several points of interest. First, the degree of agreement between the Estimation and Telegraph methods of rating importance was quite impressive ($r = .78$), especially considering that the Estimation and Telegraph ratings were also made by different subjects. Second, it appeared as though words which occurred later in a sentence were rated as more important than words which occurred earlier in a sentence. One would think that the words occurring earlier in a sentence would limit the set of possible later-occurring words and thus render them less important. Third, the two measures of importance correlated with the word-length/position-in-word/targets-in-word factor; the longer words were rated as more important. And finally, the importance and word length factors showed a strong negative correlation with word frequency and an even stronger negative correlation with log word frequency.

The number of errors made when each of the 260 target occurrences occurred in prose and in scrambled prose was obtained, and from these numbers the total errors and PSEs for each of the target occurrences was computed. The coefficients of correlation between these four measures and the 13 word variables can be seen in Table 13. Clearly, most of the large correlations in Table 13 involve the seven intercorrelated variables (estimation rating, telegraph rating, word

Table 12

Correlation Matrix for 13 Word Variables. A Correlation Coefficient of .13 Is Necessary for $p < .05$, .17 for $p < .01$, and .22 for $p < .001$, Two-tailed (df = 258)

Variable[1]	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1	-.08	.01	.23	.97	.01	.01	.04	.06	0	.06	0	.01
2	-.08	1	.16	.01	.06	-.06	.13	.07	.05	.01	.04	-.02	-.03
3	.01	.16	1	.03	0	-.22	.49	.66	.24	.41	.22	-.28	-.46
4	.23	.01	.03	1	.24	-.08	-.06	.09	.30	.25	.03	-.05	-.12
5	.97	.06	0	.24	1	-.01	.05	.03	.08	0	-.01	-.02	.01
6	.01	-.06	-.22	-.08	-.01	1	-.06	-.16	-.13	-.18	-.18	.19	.22
7	.01	.13	.49	-.06	.05	-.06	1	.27	.12	.21	.04	-.15	-.24
8	.04	.07	.66	.09	.03	-.16	.27	1	.31	.58	.14	-.44	-.76
9	.06	.05	.24	.30	.08	-.13	.12	.31	1	.78	.02	-.47	-.46
10	0	.01	.41	.25	0	-.18	.21	.58	.78	1	.19	-.58	-.75
11	.06	.04	.22	.03	-.01	-.18	.04	.14	.02	.19	1	-.10	-.22
12	0	-.02	-.28	-.05	-.02	.19	-.15	-.44	-.47	-.58	-.10	1	.64
13	.01	-.03	-.46	-.12	.01	.22	-.24	-.76	-.46	-.75	-.22	.64	1

[1] 1 = line on page, 2 = column on page, 3 = position in word, 4 = position in sentence, 5 = position in passage; 6 = pronunciation; 7 = targets in word; 8 = word length, 9 = telegraph rating, 10 = estimation rating, 11 = letter frequency, 12 = word frequency, 13 = log word frequency

Table 13

Correlation Coefficients of Errors with Word Variables. A Coefficient of .13 Is Necessary for $p < .05$, .17 for $p < .01$, and .22 for $p < .001$ (df = 258)

	Total Errors		Errors in Prose	Errors in Scrambled Prose	PSEs	
	All Subjects	Fast Readers	All Subjects	All Subjects	All Subjects	Fast Readers
line on page	.11	.13	.11	.11	.01	.12
col. on page	.01	.01	0	.02	-.03	.03
posn. in word	-.02	.02	-.05	.02	-.14	-.06
posn. in sent.	0	-.02	-.01	.01	-.03	-.03
posn. in pass.	.09	.12	.09	.08	.03	.11
pronunciation	.03	-.02	.03	.02	.01	.05
targets in word	.09	.15	.10	.08	.03	.06
word length	-.26	-.23	-.30	-.19	-.23	-.19
teleg. rating	-.23	-.20	-.25	-.19	-.15	-.13
estim. rating	-.40	-.37	-.43	-.33	-.23	-.22
letter frequency	-.14	-.14	-.13	-.13	-.02	-.04
word frequency	.53	.42	.53	.47	.18	.22
log word frequency	.41	.38	.44	.34	.23	.21

frequency, log word frequency, word length, position in word, and targets in word) which will from here on be termed the "importance/frequency factor." In fact, when the seven variables of the importance/frequency factor were entered first into the four regression equations for prediction of total errors, errors in prose, errors in scrambled prose, and PSEs, the addition of any of the other six variables to the equations failed to account for a significant amount of the remaining variation. The importance/frequency factor accounted for 36.7% of the variation of total errors, 38.5% of the variation of errors in prose, 29.1% of the variation of errors in scrambled prose, but only 7.1% of the variation of PSEs.

Since the occurrence of a negative correlation between PSEs and word importance was a major prediction of the importance hypothesis, the -.23 correlation of PSEs with Estimation ratings was subjected to further analysis. One covariate of PSEs which could have inflated the negative correlation between PSEs and Estimation ratings was total errors. In other words, the tendency to make more PSEs on target occurrences in unimportant words might have resulted entirely from the tendency to make more errors overall on the less important words. However, this was not the case, since the correlation coefficient between PSEs and Estimation ratings with total errors partialled out was still significant ($r = -.19$, $t(257) = 3.05$, $p < .01$).

The 13 word variables were also correlated with the total errors and PSEs of the 43 fast readers (who showed more PSEs than the other subjects). The correlation coefficients obtained were generally comparable to those obtained using the entire subject sample, except perhaps for a tendency in the 43 fast readers for PSEs to be positively correlated with the line on the page and position in the passage of the target occurrence (see Table 13). This raises the possibility that the additional PSEs made by those subjects were made in the later parts of the passage.

While both errors in prose and errors in scrambled prose are correlated with most of the variables in the importance/frequency factor, the correlations are stronger for errors in prose than for errors in scrambled prose. Using a test for comparison of nonindependent correlation coefficients (Ferguson, 1971), errors in prose were more highly correlated than errors in scrambled prose with word length ($t(257) = 3.39$, $p < .001$), estimation ratings ($t(257) = 3.22$, $p < .005$), word frequency ($t(257) = 2.07$, $p < .05$), and log word frequency ($t(257) = 3.29$, $p < .005$). To further investigate this effect of passage meaningfulness, the percentage of errors each subject made on the 65 highest frequency target occurrences (i.e., those in the highest frequency words) and 65 lowest frequency target occurrences was computed for both when the targets occurred in a prose passage and when they occurred in a scrambled prose passage (see Table 14). An

Table 14

Mean Percent Errors Made on High and Low Frequency Target Occurrences when Embedded in Prose and Scrambled Prose Passages

	High frequency target occurrences	Low frequency target occurrences
Prose	15.98	3.78
Scrambled Prose	12.48	4.25

analysis of variance performed on the arcsin transformation of these percentages indicated that the high frequency target occurrences were more often missed than low frequency ones ($F(1,199) = 242.77$, $p < .001$), and that more errors were made when these high and low frequency target occurrences were surrounded by a prose context ($F(1,199) = 8.92$, $p < .005$). Further, the effect of the frequency of the target-containing words was larger when they were embedded in prose than when they were embedded in a nonprose context ($F(1,199) = 16.35$, $p < .001$). This confirms the correlational differences mentioned above, and suggests that importance/frequency is more likely to affect the probability of a target occurrence being missed when the target occurs in prose than when it occurs in a scrambled prose context.

Effects of Serial Position and Passage Type

To assess if the serial position of a passage in the booklet affected the total number of errors made on the target occurrences in that passage, a one-way analysis of variance was run on the mean number of errors for each passage position. The differences between these means were not significant whether subjects was used as the random effects variable ($F(9,1791) = 1.09$, $p > .20$) or passage type was used as the random effects variable ($F(9,81) < 1$).

There was an effect of serial position of the passage on the mean number of PSEs ($F(9,1980) = 2.42$, $p < .01$), but this

effect did not appear to be reliable across passage type, since it was not significant when passage type rather than subjects was used as the random effects variable ($F(9,81) = 1.74$, $p < .10$). Moreover, while the pattern of the effect of serial position on recognition scores ($F(9,1791) = 7.10$, $p < .001$) corresponded to at least the recency part of the bowed serial position curve typically found in memory experiments (see Figure 2), the pattern of effects of serial position on PSE (Figure 2) was so irregular as to be uninterpretable.

The mean number of total errors and PSEs per passage per subject for each of the ten passage types can be seen in Table 15. Separate analyses of variance showed that total errors varied among the passage types ($F(9,1791) = 94.04$, $p < .001$), but that the effect of passage type on PSEs was not significant ($F(9,1980) < 1$). However, it is possible that this latter effect is not statistically significant because the pattern of the subject-passage assignments forced the analysis of variance to be done using a completely randomized design, rather than the more powerful repeated measures design used to test the effects of passage type on total errors.

Since the importance/frequency factor is a major predictor of the probability of an error in a given word, it is possible that some of the ten passage types showed more search errors than others because they just happened to contain many target words of low importance/high frequency. To

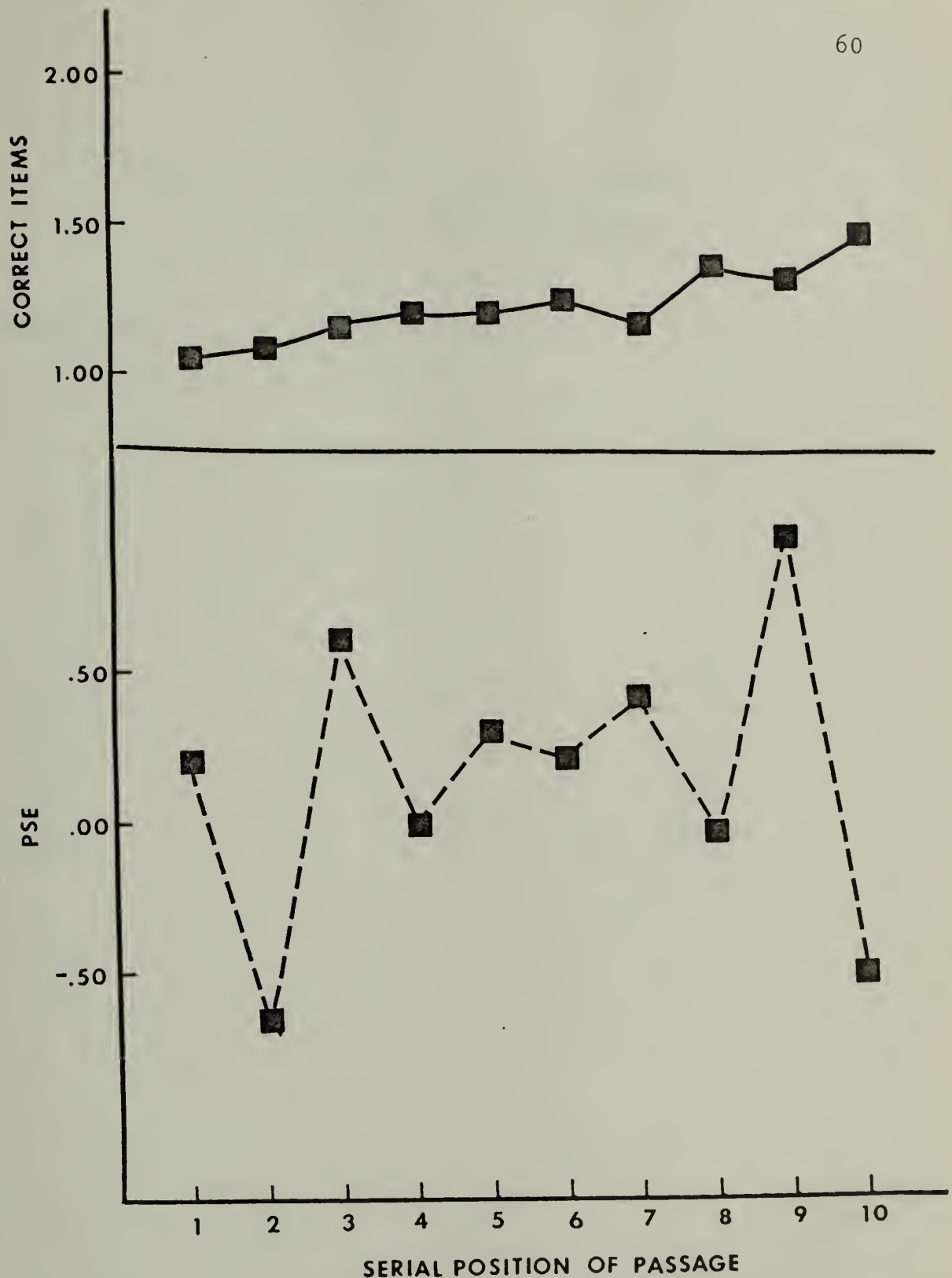


Figure 2. Mean recognition scores per passage per subject, and mean PSEs per passage per subject as a function of serial position of the passage in the booklet.

Table 15

Mean Total Errors and PSEs per Passage per Subject for Each Passage Tyle

Passage Type	Total Errors	PSEs
A	1.42	.15
E	1.70	.10
I	.77	-.20
O	.65	.07
F	2.33	.01
H	4.05	-.09
N	3.26	-.05
S	1.21	.30
T	1.23	.58
W	1.18	.59

gauge the extent to which such differences accounted for the different numbers of errors made on each passage type, one mean value of each of the seven variables in the importance/frequency factor was computed for each of the ten passage types. A multiple regression indicated that those seven variables accounted for 83.5% of the variation in total errors between passages (and 60.4% of the between-passage variation in PSEs), thus supporting the idea that differences in the importance/frequency of the target-containing words of a passage were largely responsible for the differences in total errors between the passages.

Commission Errors

As mentioned earlier, 15 of the visual search errors made in the present study consisted of the subject circling a letter other than the target letter for the passage. The particular letters that were mistaken for the correct target letters can be seen in Table 16. Inspection of Table 16 indicates that most of the substituted letters are similar to the correct letters on visual dimensions, but not on acoustic dimensions. On every one of the nine occasions where the correct target letter was an "ascending" letter, the substituted letter was also an ascending letter. Other types of visual confusions were also represented: m was circled in place of its inverted form w, and v was substituted for w which is formed by the combination of two vs. In general,

Table 16

Commission Errors Made by the 200 Subjects

	Subject Number	Correct target letter	Substituted target letter
Read instructions	A-35	f	l
	A-39	h	f
	B-15	w	m
	B-18	e	a
Search instructions	C-2	s	a
	C-4	h	d
		i (or I)	l
	C-19	h	f
	C-24	h	f
	C-38	e	No
	C-40	t	i
	C-43	w	b
	D-2	f	h
	D-17	t	l
		t	l

these commission errors provide evidence that the search in this visual search task actually was visual.

C H A P T E R I V

DISCUSSION

Effects of Passage Meaningfulness

The present results fail to provide conclusive evidence that the meaningfulness of the arrangement of words in a passage generally affects the number or distribution of visual search omission errors. There was no significant difference between the number of errors occurring in prose passages and the number occurring in the scrambled prose passages; and, although the pattern of errors in certain portions of certain passages appeared to be affected by passage meaningfulness, the present results do not indicate that this is a generally occurring effect. Although the number of PSEs which did occur was significantly correlated with at least one measure of word importance, this correlation was relatively small compared with the correlation between the same measure of word importance and the total number of errors at a target position.

Although the present results indicate that if there is an effect of passages meaningfulness on visual search omission errors, it is a small one, it is possible that weaknesses in the design of the present study are responsible for the failure to observe the effect. First, the procedure of giving subjects one minute to complete all passages probably caused artifactual effects of passage meaningfulness on error

distribution, and reduced the ability of the study to reveal passage meaningfulness effects caused by automatic perceptual processes. Second, the relatively low error rate found in this study made the differences between all conditions smaller and therefore more difficult to detect through the haze of random variability. If the visual search task had been conducted as a reaction time task, both problems would have been minimized. Subjects who often required more than one minute would be able to take the additional time, so they would not have to work more rapidly at the end of some passages and not others. And, the subjects who usually finished in less than a minute would not be able to check their work, and thus the overall error rate would be higher.

The reason that the study was not originally designed as a reaction time task is because it would have then been possible to argue that any PSEs found resulted from a speed-accuracy trade-off. However, if the pattern of results was not changed by instructions, it could be argued that rapid perceptual processing is an automatic response to meaningful word sequences and is thus interesting. Moreover, equalizing the time subjects spent on the prose and scrambled prose passages does not really circumvent the speed-accuracy trade-off argument, since, although many subjects were observed checking their work, that is by no means any assurance that all subjects spent equal amounts of time searching their prose and scrambled prose passages.

A third weakness in the design of the present study was that of having each subject search the scrambled prose passages which were not formed from the prose passages that he searched. This made meaningfulness a between-subjects variable for each passage and added variability to the overall error rates for prose and scrambled prose. This was done in order to prevent the possibility of subjects increasing their accuracy on later passages by remembering the exact positions of the target letters from an earlier passage of the same passage type. While this would be a real danger of the totally within-subjects design, Healy's (personal communication) failure to find such effects and the large amount of between-subjects variability shown in this task make the totally within-subjects design seem an attractive alternative to the design of the present study.

The failure to find a difference between the number of errors made in prose and scrambled prose passages contradicts both Healy's (1976) and Schindler and Jacobs' (1976) finding of more errors in prose than in scrambled prose. Healy's effect was relatively small, and so might have been due to her use of reaction time instructions and a within-subjects design. Schindler and Jacobs' effect could have been due to the fact that in their study, a subject searched either three sentences or three scrambled sentences. If PSEs are produced by voluntary strategies, this procedure may have exaggerated the effect. However, the present study's failure to find

that more PSEs occurred on the first passage of a booklet than on later ones (see Figure 2) renders this explanation unlikely. A more likely reason for the meaningfulness effects found by Schindler and Jacobs in the F- and N-sentences (see Figure 1) was that they were simply due to some property peculiar to those sentences. The fact that the F-sentence was selected precisely for its impressive ability to cause people to miss the fs in of (e.g., O'Neill & Ruder, 1974, p. 91) raises the probability that it (and the similarly-structured N-sentence) could have had some property not shared by sentences in general.

Effect of Word Frequency

The results indicated that the seven intercorrelated variables of the importance/frequency factor account for most of the total error and PSE variability that was accounted for by any of the word variables studied. Of the seven variables, word length, word frequency, and Estimation ratings tended to have the highest correlation coefficients with total errors and PSEs. Is one of these variables the basic variable and the others merely correlates? Healy (1976, Expt. 4) controlled for word length and still found an effect of word frequency, thus implicating word frequency as the more fundamental of the two variables. One would think that it is rather difficult to directly estimate the importance of words in a prose passage. Since it has been shown (e.g., Carroll,

1971) that subjects can fairly accurately estimate the frequency of a word, it is possible that the subjects in this task simply rated the more frequent words as less important. In fact, using a rating method similar to the Estimation rating method of the present study, Galbraith and Underwood (1973) found that the correlation coefficient between subjective ratings of frequency and Kucera-Francis frequency equalled .63; coincidentally, the correlation coefficient between Estimation ratings of importance and Kucera-Francis frequency found in the present study equalled -.58.

Moreover, it is probable that the degree of correlation between word frequency and errors found in the present study would have been much higher if the morphemes in multi-morpheme words had been considered as separate items. For instance, the es in the word overuse were missed very frequently, yet the Kucera-Francis frequency of overuse is 0; and, the second n in depending was missed by 88 of the 200 subjects, yet the Kucera-Francis frequency of depending is only 32 occurrences/million. Thus, let us say, at least tentatively, that word frequency is the potent variable of the seven variables in the importance/frequency factor.

It is worth noting which variables did not appear to affect visual search omission errors. Position on the page or position in the passages did not appear to have much effect, except perhaps for the tendency for the fast readers to make more errors and more PSEs toward the later-occurring target

letters in a passage. Position in the sentence did not have any effect, or at least no linear one. The absence of significant correlations between pronounceability and errors is evidence against the use of at least one particular acoustic scanning strategy. And, the preponderance of visual confusions in the commission errors is further evidence against the importance of acoustic factors. However, since 44% of the subjects did report using the sound of the target letter at least 25% of the time, the use of acoustic factors cannot be ruled out by the present results. In any event, the demonstration of the high correlation between word frequency and errors in prose raises questions about Corcoran's (1966) finding that unpronounced es are missed more often than pronounced es, since he did not control for the frequency of the target-containing words.

The correlation of word frequency with the number of errors made in the scrambled prose passages was large and highly significant. This result, in combination with the failure to find large effects of passage meaningfulness, suggest that Schindler and Jacob's importance hypothesis is, in fact, not correct. If there are any units larger than words, in prose, they must be so weak as to be insignificant to the visual processes of most subjects. Rather, the high correlations of both errors in prose and errors in scrambled prose with word frequency suggest that Healy's view that frequent words are read as units is a more accurate way of describing

the situation.

In fact, since the unitization hypothesis holds that frequent words act as units during reading, the passage meaningfulness effects which were found in the present study may be explainable by differences in the tendency of the subjects to read. Both recognition test scores and questionnaire data indicated that subjects were more likely to read the words in prose passages than in scrambled prose passages. So, if frequent words act as units only when reading, it would be expected that there would be a larger difference in the number of errors made in high and low frequency words embedded in prose than in those embedded in scrambled prose. And, if words acting as units causes a tendency to miss the second target letter in a word, it would be expected that this tendency would be more pronounced when words are embedded in prose than when they are embedded in scrambled prose. Moreover, recognition score data (see Table 10) provide at least weak evidence that the larger number of PSEs shown by the fast readers is due to their tendency to be more likely to read the prose passages and less likely to read the scrambled prose passages than were the average and slow readers.

Thus, the unitization hypothesis appears to account for the present results fairly well. Not only did high frequency words conceal their letters whether they were embedded in prose or scrambled prose, but also those passage meaningful-

ness effects which were found may be explainable by the tendency of subjects to be more likely to read the prose passages than the scrambled prose passages and an increase in this tendency among fast readers. However, the unitization hypothesis leaves unexplained why this tendency exists. It could be argued that reading is simply a voluntary strategy for searching word sequences and it is used whenever the subject notices he can save time by reading. But in the present experiment, the instructions stressed accuracy rather than speed, and most subjects were given more time than they needed to search through the passages once. In fact, instructions which explicitly told the subjects not to read and told them that reading would impair accuracy in the task appeared to have no effect on most subjects. Thus, the possibility arises that the subjects' tendency to more often read when searching prose than when searching scrambled prose may be due to a strategy shift which is not completely voluntary.

Modes of Visual Processing

The word "strategy" implies a voluntary sequence of processes. There are two senses in which the visual processes used in reading are not completely voluntary. First, a great deal of practice is necessary before one acquires the capacity to use the visual processes of reading. You cannot use a reading strategy if you haven't practiced it; in other words, it is a skill. Second, to an extent, it may not be

completely a matter of conscious choice whether the visual processes of the reading skill are used. In fact, it may be a property of all skills that the more skilled you become, the more obligated you are to use the skill when dealing with the stimuli to which the skill provides practiced responses. For example, a highly skilled tennis player may find it hard to hit an incoming ball incorrectly, even if it is necessary to do so in order to illustrate an error to a student.

Thus, the visual processes of reading, which have been observed in this study, may be more appropriately termed the "prose mode" rather than the "reading strategy." The present results, as well as those of Healy (1976), suggest that words arranged to have the visual features of prose are sufficient to elicit the processes of the prose mode, although the presence of the syntactic and semantic properties of prose can serve to increase the probability that the prose mode is elicited. Probably, every visual feature of normal prose which, when altered, makes reading more difficult, plays a role in eliciting the prose mode. One reason for Krueger's (1970) finding that subjects make fewer errors in prose than in scrambled prose may have been his presentation of the passages in all uppercase letters, since prose printed in uppercase letters is read more slowly than prose printed in mostly lowercase letters (Tinker, 1955).

If there is a prose mode, there may also be a separate mode for the visual processing of symbol strings which occur

outside of prose. The processes behind this skill may comprise what could be called a "word mode." There may also be a "letter mode" which may be responsible for the rapid recognition of all familiar shapes, including, of course, those which comprise the letters of the alphabet.

Postulating the existence of separate processing modes for different kinds of visual stimuli offers the possibility of explaining some of the contradictory findings on the effects of familiarity on visual processing. The present results indicate that high frequency words conceal their letters when a prose-like array is being searched, but Krueger (1970, Expt. 3) found that high frequency words reveal their letters when they are presented in a two-word array. Many studies have confirmed Krueger's finding that isolated words reveal their letters (e.g., Eichelman, 1970; Reicher, 1969), but it has not been found that words reveal visual features of letters (Earhard & Fullerton, 1969; Cohen, 1974), and words may even conceal such features (Pillsbury, 1897; Postman, Bruner, & Walk, 1951). However, it has been found that isolated letters reveal their features (Ambler & Proctor, 1976).

These findings can be simplified if it is assumed that a mode of visual processing consists of information concerning how relevant visual units are likely to be combined. This information enables later-occurring units to be identified more rapidly since some information about their identity is

provided by the earlier-occurring units and thus less stimulus information need be processed. If the units of the prose mode are words (or some other multi-letter form), the the faster transition from word to word would facilitate search for a word. However, since these more rapid transitions involve less stimulus processing, the prose mode would make search for letters and features of words more difficult. If the units of the word mode are letters, then visual processing in the word mode should facilitate search for letters, but interfere with search for subletter features. By the same reasoning, the visual processes of the letter mode may be expected to facilitate search for lines and curves but may interfere with search for irrelevant features such as thickness of lines.

This view that there exists separate prose, word, and letter modes of visual processing is, of course, only hypothetical and will require a great deal of empirical verification. Further, although it may simplify the study of letter, word, and prose perception, it does not simplify the study of perception in general, since there may well be as many modes as there are visual tasks. However, this view does have some use in that it can serve to constrain and direct further theorizing.

Finally, how does the automatization hypothesis stand in the light of all of this? Unfortunately, the present results have not provided new evidence to directly support

it, since no convincing evidence was found for even primitive multiword units. However, this may simply mean that the specific word sequences of prose are not familiar enough for multiword units to be apparent. Moreover, the view of perceptual processing which is suggested by the present results is consistent with the automatization hypothesis, and shares with it the emphasis on the role played by repeated tasks in determining the form of the perceptual representations which we acquire.

REFERENCES

Ambler, B. A. & Proctor, J. D. The familiarity effect for single-letter pairs. Journal of Experiment Psychology: Human Perception and Performance, 1976, 2, 222-234.

Barnhart, C. L. (Ed.) American college dictionary. New York: Random House, 1966.

Carroll, J. B. Measurement properties of subjective magnitude estimates of word frequency. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 722-729.

Cohen, R. L. The perception of mutilated letters in words and nonwords. Unpublished manuscript, University of Massachusetts, 1974.

Corcoran, D. W. J. An acoustic factor in letter cancelation. Nature, 1966, 210, 658.

Corcoran, D. W. J. Acoustic factor in proof reading. Nature, 1967, 214, 851-852.

Corcoran, D. W. J. & Weening, M. Acoustic factors in visual search. Quarterly Journal of Experimental Psychology, 1968, 20, 83-85.

Craik, F. I. M. & Lockhart, R. S. Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 671-684.

Crosland, H. R. An investigation of proofreaders' illusions. University of Oregon Publication, 1924, 2(6).

Dewey, G. Relative frequency of English spellings. New York: Teachers College Press, 1970.

Earhard, B. & Fullerton, R. How much does repetition facilitate perception? Journal of Experimental Psychology, 1969, 81, 101-108.

Eichelman, W. H. Familiarity effects in the simultaneous matching task. Journal of Experimental Psychology, 1970, 86, 275-282.

Furguson, G. A. Statistical analysis in psychology and education. New York: McGraw-Hill, 1971.

Galbraith, R. C. & Underwood, B. J. Perceived frequency of concrete and abstract words. Memory and Cognition, 1973, 1, 56-60.

Glass, G. V., Peckham, P. D., & Sanders, J. R. Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. Review of Educational Research, 1972, 42, 237-288.

Healy, A. F. Detection errors in the word the: Evidence for reading units larger than letters. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 235-242.

Howes, D. & Solomon, R. L. Visual duration threshold as a function of word probability. Journal of Experimental Psychology, 1951, 41, 401-410.

Kreuger, L. E. Search time in a redundant visual display. Unpublished doctoral dissertation, Harvard University,

1969.

Kreuger, L. E. Search time in a redundant visual display. Journal of Experimental Psychology, 1970, 83, 391-399.

Kucera, H. & Francis, W. N. Computational analysis of present-day American English. Providence, R.I.: Brown University Press, 1967.

Miller, G. A., Galanter, E., & Pribram, K. H. Plans and the structure of behavior. New York: Holt, Rinehart, and Winston, 1960.

O'Neill, C. & Ruder, A. The complete guide to editorial freelancing. New York: Dodd, Mead & Co., 1974.

Piaget, J. The mechanisms of perception. New York: Basic Books, 1969.

Pillsbury, W. B. A study in apperception. American Journal of Psychology, 1897, 8, 315-393.

Postman, L., Bruner, J. S., & Walk, R. D. The perception of error. British Journal of Psychology, 1951, 42, 1-10.

Reicher, G. M. Perceptual recognition as a function of meaningfulness of stimulus material. Journal of Experimental Psychology, 1969, 81, 275-280.

Schindler, R. M. & Jacobs, P. I. What do we see when we read? Paper presented at the Annual Meeting of the Eastern Psychological Association, New York City, 1976.

Tinker, M. A. Prolonged reading tasks in visual research. Journal of Applied Psychology, 1955, 39, 444-446.

Winer, B. J. Statistical principles in experimental design (2nd ed.) New York: McGraw-Hill, 1971.

APPENDICES

A. The 20 passages used in the study

B. Examples of a telegraph rating sheet and an Estimation rating sheet

C. A copy of the word recognition test

D. The questionnaire used in the study

E. The instructions used in the study

APPENDIX A

A, a

Something strained within me started to relax when I began to approach the falls. I heard the distant music of the spring water which had been running constantly for uncounted years. Soon, I was swept away with a sense of permanence, of freedom, and of extreme beauty. It appeared to me that the forest was such a firm, dependable thing.

A,a

Uncounted strained been when started it relax such of began me approach for falls. To heard I distant swept extreme the I water sense had something of constantly the within of years. The, to was thing away firm a which me permanence, forest the, and soon music beauty. To appeared spring that I running was of a with, dependable freedom.

E,e

No society is so rich that it can afford to waste its educated brain power. We cannot continue to underuse our trained young women and after that overuse and casually expend our trained young men. Given man's greater inborn susceptibility to life's many stresses, we must stop compounding his problems by insisting he carry society's sad burdens on his shoulders.

E,e

It society brain compounding insisting rich waste can educated carry power. We inborn continue no underuse and trained man's women its after stop overuse his that our expend sad trained to is men. Given young greater cannot susceptibility so life's must stresses, we to young casually his and problems afford by to he many society's that burdens our on shoulders.

I,i

Despite death being the surest of certainties, modern society seems to be keeping a conspiracy of silence about it. Recently, I addressed the presidents of several prestigious colleges concerning the missing elements in college education. I made the point that while we teach people how to make a living, we fail to teach them how to face life's stern realities.

I,i

Despite seems being colleges that certainties, the to society a addressed keeping to conspiracy we silence death it. Elements, I surest modern presidents about make prestigious recently concerning how missing them the in several education. I people a point face while stern college the teach we of living, be fail of made to how the to of life's teach realities.

O,o

Here, the peace and quiet was almost tangible. Depending on the season, I might linger to inhale the soft fragrance of the grape blossoms, feast on ripening grapes, or peer into a bird's nest holding bright eggs or fuzzy heads with big mouths. Busy ants moved in close ranks across the wooden ramps, intent upon a mysterious, but vital goal.

O,o

Fragrance, ants bird's grape almost a inhale. With here on and season, intent the was to a might but soft depending of big the I blossoms, vital on linger grapes, or eggs into peer ripening holding the bright or ramps heads tangible mouths. Nest the moved busy close fuzzy across feast wooden peace, ranks upon in mysterious, quiet the goal.

F,f

Even after several days of captivity, the wild eagle refused the food I offered it at the end of a stick. Finally, on the fifth day, it took a piece of meat, fearfully but gently, from my fingers. In fact, before long, it had become so comfortable that it would sometimes fly away from the cage and return without prodding.

F,f

Cage after without away of sometimes, prodding it my refused and food a offered meat the took of I eagle. Finally, it day fifth the, captivity in at of long, fearfully the return, from it fingers. So fact, before even, end several that comfortable wild become stick the it fly days from gently a the had but piece on would.

H,h

The vulnerability of the marine environment becomes clear when we consider that even though oceans blanket over half of the entire earth, their productivity is limited wholly to the extremely rich waters over the continental shelves. Most of the world's fish are harvested from these usually shallow waters, which make up only a slight fraction of the total sea area.

H,h

The consider becomes the usually clear make vulnerability when environment that from though productivity marine half we the waters earth, their of continental fraction wholly is the over even rich blanket sea the over entire shelves. Limited the only of fish up harvested area these total a of shallow of are, which extremely most slight oceans to the waters world's.

N,n

Often the brain or other organs risk being deprived of important sources of energy by inadequate diets. The brain, for instance, depending largely on the blood's nutrients for its energy, may function somewhat abnormally with certain diets. Insufficiently large quantities of carbohydrates and fats found mostly in the proteins of meat and fish can result in mild anxiety or depression.

N,n

Often may brain of diets organs fats being other the or important the mostly energy or inadequate large. For brain, its instance, depending blood's on of somewhat nutrients risk by energy, for function deprived abnormally fish certain with. Insufficiently diets quantities carbohydrates of and mild found the in result proteins largely and the can sources in of anxiety meat depression.

S,s

Everybody seems to assume at the start that labor is entitled to an increase in wages, regardless of what is happening to the economy. Nobody argues anymore that increases in productivity should be passed on in the form of price reductions. Business often assumes that it must yield wage increases that are at least equal to cost of living increases.

S,s

Happening seems at assume living start anymore to is are wage often increase be wages, regardless economy is at it productivity price. An the argues form nobody increases labor everybody should in passed yield the equal entitled reductions. Business in of assumes what of must that of to increases the that to least that to cost that in on increases.

T,t

The mild nature of a cheetah makes this jungle cat a most easily domesticated wild animal. There is no record of an unprovoked attack on man. In fact, once tamed, a cheetah seems to even take pleasure in pleasing. He learns quickly, can be taught how to find and retrieve wooden sticks, and eventually becomes totally loyal to his master.

T,t

The find nature even cheetah loyal this learns cat no most jungle domesticated pleasing he. There of animal unprovoked an seems attack a once. Be fact, a is tamed, in cheetah of an to mild take becomes and can wild. Pleasure wooden a, record taught and to in makes retrieve easily sticks, his eventually quickly totally man to how master.

W,w

I awoke early the next morning, and slowly made my way over toward the window. The dawn sun was quite low over the water, and the wind still was surprisingly brisk. A white gull slid like a shadow between the jagged rocks which lined the beach. The sea was again filled with heavy waves, and I knew I was stranded.

W,w

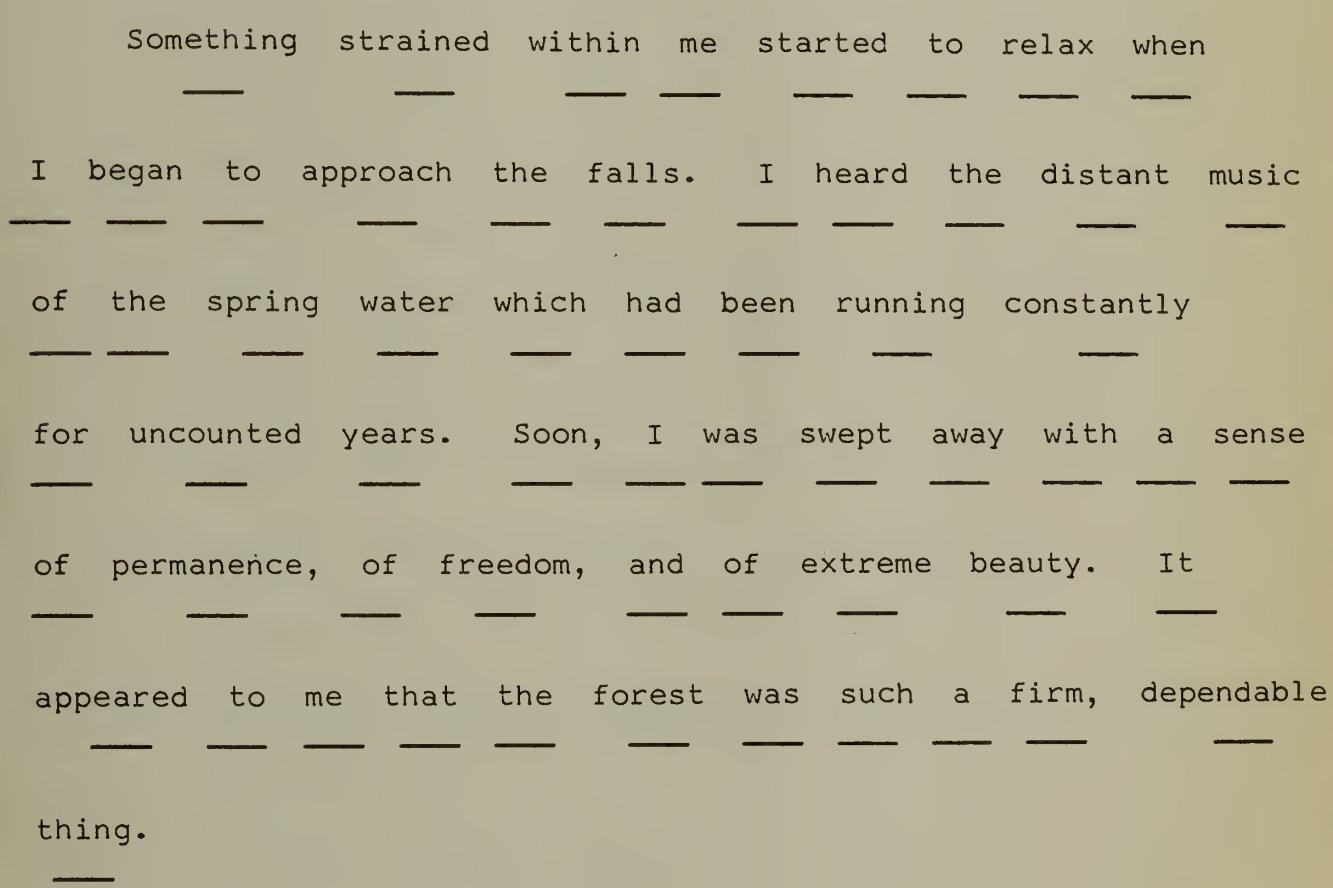
A awoke gull next stranded my, the slowly morning way like toward and window. And dawn the was beach low filled I water, the and wind rocks was the slid sun again. I white surprisingly made shadow between sea the over the which lined brisk the. I heavy was early jagged with quite waves, still knew a was the over.

APPENDIX B

Something strained within me started to relax when I began to approach the falls. I heard the distant music of the spring water which had been running constantly for uncounted years. Soon, I was swept away with a sense of permanence, of freedom, and of extreme beauty. It appeared to me that the forest was such a firm, dependable thing.

Something strained within me started to relax when I began to approach the falls. I heard the distant music of the spring water which had been running constantly for uncounted years. Soon, I was swept away with a sense of permanence, of freedom, and of extreme beauty. It appeared to me that the forest was such a firm, dependable thing.

Something strained within me started to relax when I began to approach the falls. I heard the distant music of the spring water which had been running constantly for uncounted years. Soon, I was swept away with a sense of permanence, of freedom, and of extreme beauty. It appeared to me that the forest was such a firm, dependable thing.



Instructions: Only one of the three words in each row has appeared in the paragraphs or scrambled paragraphs you have just seen. In each row, circle that word. Please do not refer to any of the other pages of this booklet.

1.	manner	spring	theory
2.	whip	fats	figs
3.	child	party	music
4.	death	court	voice
5.	night	group	young
6.	diets	rhyme	quart
7.	grapes	brandy	castle
8.	issue	labor	stand
9.	swamp	photo	eagle
10.	price	visit	scene
11.	myth	knee	eggs
12.	cork	aide	cage
13.	fray	gull	plum
14.	marine	motion	device
15.	pure	wild	fort
16.	magnet	oceans	eskimo
17.	treaty	hotels	jungle
18.	teach	smoke	bread
19.	beach	home	bible
20.	sequel	inborn	gunner

APPENDIX D

1. Age _____

2. Sex: _____ Male _____ Female

3. Dominant hand (the hand you write with):

   _____ Left _____ Right

4. Year: _____ Freshman

   _____ Sophomore

   _____ Junior

   _____ Senior

   _____ Graduate student

5. Approximately how many pages of printed material (i.e., books, magazines, newspapers) do you read during an average week? (Exclude material required for courses.)

   _____ 0-25 pages

   _____ 25-50 pages

   _____ 50-100 pages

   _____ 100-200 pages

   _____ more than 200 pages

6. How would you guess that you reading speed compares with that of other college student? Your reading speed is

   _____ much faster

   _____ above average

   _____ average

   _____ below average

   _____ much slower

1.(a) How often did you find yourself reading as you were circling the occurrences of the target letters in the paragraphs?

_____ less than 5% of the time

_____ 5%-25% of the time

_____ 25%-75% of the time

_____ 75%-95% of the time

_____ more than 95% of the time

(b) How often did you find yourself reading as you were circling the occurrences of the target letters in the scrambled paragraphs?

_____ less than 5% of the time

_____ 5%-25% of the time

_____ 25%-75% of the time

_____ 75%-95% of the time

_____ more than 95% of the time

2. How often did you find yourself using the sound (i.e., pronunciation) of the target letter when circling each occurrence of it?

_____ less than 5% of the time

_____ 5%-25% of the time

_____ 25%-75% of the time

_____ 75%-95% of the time

_____ more than 95% of the time

3. You circled all occurrences of target letters in paragraphs and in scrambled paragraphs. Which did you find more difficult?

_____ paragraphs

_____ scrambled paragraphs

_____ no difference

APPENDIX E

READ instructions:

The following is an experiment designed to study the visual processing of printed material. As with all psychology experiments, you are free to withdraw from participation at any time.

The top page of the booklet I've handed out to you is a short questionnaire. Please fill out this questionnaire.

On each of the next ten pages of the booklet is a sequence of words. Some of the word sequences form paragraphs, but others result from scrambling the words of a paragraph to form nonsense. Above each of the word sequences is the target letter for that page. There will be a different target letter for each page. What I would like you to do is to read the sequence of words and circle the target letter every time it occurs. For example, if the target letter is "A", circle every "A" which occurs in the words on the page. On each page, the target letter is in both upper- and lowercase form. This is to remind you that it makes no difference whether the target letter is upper- or lowercase: Circle every single one.

You will have one minute to do each page. If you finish before time is up, go back and check your work. Please be as accurate as possible. When I say "Start," turn to the next page and begin. When I say "Stop," put down your pencil, and take a brief rest.

Any questions?

SEARCH instructions:

The following is an experiment designed to study the visual processing of printed material. As with all psychology experiments, you are free to withdraw from participation at any time.

The top page of the booklet I've handed out to you is a short questionnaire. Please fill out this questionnaire.

On each of the next ten pages of the booklet is a sequence of words. Some of the word sequences form paragraphs, but others result from scrambling the words of a paragraph to form nonsense. Above each of the word sequences is the target letter for that page. There will be a different target letter for each page. What I would like you to do is to search through the sequence of words and circle the target letter every time it occurs. For example, if the target letter is "A", circle every "A" which occurs in the words on the page. On each page, the target letter is in both upper- and lowercase form. This is to remind you that it makes no difference whether the target letter is upper- or lowercase: Circle every single one.

I would like you to search through the word sequences from left to right and top to bottom as <u>if</u> you were reading, but under <u>no</u> circumstances should you <u>read</u> the words. Reading will cause you to miss instances of the target letter. Rather than read, you should search each word letter by letter, and totally ignore its meaning.

You will have one minute to do each page. Please be as accurate as possible. If you finish before time is up, go back and check your work. Remember, accuracy is of the utmost importance. When I say "Start," turn to the next page and begin. When I say "Stop," put down your pencil and take a brief rest.

Any questions?

RECOGNITION TASK instructions:

Please turn to the next page of the booklet. On this page, there are 20 rows of three words. One and only one of the three words in each row has appeared in the paragraphs or scrambled paragraphs you have just seen. In each row, circle that word.

Please do not refer to any of the other pages of this booklet. If you are unsure of which one of the three words has appeared in the paragraphs, please guess; do not leave any rows blank. Thus, you should circle one and only one word in each of the 20 rows.

Any questions?

SECOND QUESTIONNAIRE:

On the next page of the booklet there is another short questionnaire. Please turn to this page and fill out the questionnaire.

TELEGRAPH TASK instructions:

On each of the next five pages of the booklet, a 60-word paragraph is printed twice.

When I tell you to start, I would like you first to read the top paragraph. Read it slowly and carefully and make sure you understand what the paragraph is trying to say.

Then I would like you to pretend that you want to send this paragraph as a telegraph message to a friend of yours. However, there are 60 words in the paragraph and you have only enough money to send a 40-word message. Thus you must leave out 20 of the words in the paragraph when you transmit it to your friend.

Since you must leave out 20 of the words in the paragraph, it makes sense to leave out those words which are least important to the meaning of the paragraph. Note that there are some words which can be left out of the paragraphs without at all hurting the chances that your friend will understand what the paragraph is trying to say.

Think about which 20 words of the paragraph you would leave out if you had to send it as a 40-word telegraph message. Then go to the paragraph at the bottom of the page and cross out (in a complete way) those words which you would leave out of the paragraph. (Note that one-letter words count as words, but that periods and commas are automatically included in the telegraph message and do not count as words.)

There are no right or wrong answers in this task. I am interested simply in which 20 words you consider to be least important to the meaning of the paragraph. Those are the 20 words which you should cross out.

You will have 5 minutes to complete each paragraph. Make sure that you have crossed out <u>exactly 20</u> of the words in the lower paragraph. As you are crossing out words, you can refer to the upper paragraph, but please do not refer to any page in the booklet other than the one you are working on.

When I say "Start," turn to the next page and begin. When I say "Stop," put down your pencil and take a brief rest.

Any questions?

ESTIMATION TASK instructions:

On each of the next five pages of the booklet, a 60-word paragraph is printed twice.

When I tell you to start, I would like you first to read the top paragraph. Read it slowly and carefully and make sure you understand what the paragraph is trying to say.

Then I would like you to rate how important you think each word in the paragraph is to the meaning of the paragraph. You can do this by using numbers to rate importance. If you think that no one could fully understand the paragraph if a certain word were left out, then give that word a "5" for importance. If you think that the meaning of the paragraph would be just as clear without a certain word, then give that word a "1" for importance. If the importance of a word is somewhere in between, give it a number somewhere between 1 and 5 (you can use decimal fractions if you like). In all cases though, the more important the word is for understanding the meaning of the paragraph, the higher a number you should give to that word.

Note that there are no right or wrong answers in this task. I am interested simply in how important you think each word is.

When thinking about how you will rate each word, please refer to the top paragraph. When you decide on your rating for a word, go to the bottom paragraph and write the number you have decided on in the space provided directly under that word. I will hand out a slip of paper that will help you to keep the importance rating scale in mind. You can refer to this slip of paper, but please do not refer to any page in the booklet other than the one you are working on.

You will have 5 minutes to rate each of the 60 words in a paragraph. I realize that this task is not easy. If you cannot decide on how to rate a particular word, then just guess a number between 1 and 5. But make sure that you have written one number below each of the 60 words in the lower paragraph.

When I say "Start," turn to the next page and begin. When I say "Stop," put down your pencil and take a brief rest.

Any questions?