1988

Stroop task latencies as a function of reading ability, context, and context-target relatedness.

Barbara A. Greene
University of Massachusetts Amherst

STROOP TASK LATENCIES AS A FUNCTION OF READING ABILITY, CONTEXT, AND CONTEXT-TARGET RELATEDNESS

A Thesis Presented

by

Barbara A. Greene

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

Master of Science

September 1988

Psychology Department

STROOP TASK LATENCIES AS A FUNCTION OF READING ABILITY, CONTEXT, AND CONTEXT-TARGET RELATEDNESS

A Thesis Presented

By

Barbara A. Greene

Approved as to style and content by:

James M. Royer, Chairperson of Committee

Charles Clifton, Member

Arnold Well, Member

Seymour Berger, Department Chair
Psychology Department

ACKNOWLEDGEMENTS

I would like to express my thanks to all the people who assisted me on this project. I would first like to thank my advisor, Mike Royer, whose greatly appreciated support and guidance was crucial. I would like to thank Arnie Well whose comments on the different stages were always supportive and inspired clear thinking. Special thanks go to Chuck Clifton who provided a tremendous amount of assistance on both the theoretical and practical ends. I would also like to extend my thanks to Stephen Callahan for getting me into the schools to collect data. Special thanks go to my research assistant, Sara Fishburn, for her help in fine-tuning the experiment and running subjects. Finally, I would like to thank all of my colleagues and friends who were kind enough to be pilot subjects.

ABSTRACT

STROOP TASK LATENCIES AS A FUNCTION OF READING ABILITY, CONTEXT, AND CONTEXT-TARGET RELATEDNESS

September 1988

Barbara A. Greene, B.A., Boston University

M.A., University of Massachusetts

Directed by: Professor James M. Royer

A modified Stroop task was used to identify sentence processing differences between 18 skilled and 15 less-skilled fifth grade readers. Based on the findings of Merrill, Sperber, & McCauley (1981) differences in color-naming latencies were expected to be observed as a function of reading ability, the context stimuli (sentence or word), and the target (color) word. Target words that were related to the context were expected to be associated with longer color-naming latencies relative to the latencies found with unrelated target words. Differences between the ability groups were expected to provide evidence that less-skilled readers encode word meanings while reading a sentence that are both related and unrelated to the sentence context while skilled readers encode only meanings that are consistent with the sentence context. The expected differences were not found. Color-naming latencies did not vary as a function of context-target relatedness for either ability group. The less-skilled readers were slower overall in color-naming. The

absence of longer latencies relative to a control was not in accordance with other research that has utilized similar modified Stroop tasks, so a follow-up study with 26 college students was conducted using the same methodology. The college students demonstrated longer latencies relative to a control only in the single word condition. It was concluded that methodological factors probably resulted in the present findings.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The purpose of the present study was to replicate the findings of Merrill, Sperber, & McCauley (1981) in preparation for a second study which could utilize their Stroop task paradigm to examine the effects of sentence processing instructions on the color-naming latencies of skilled and less-skilled readers. Rationales for both studies will be presented, but the major focus of the thesis will be the replication.

Merrill et al. (1981) compared interference effects for skilled and less-skilled readers on a modified Stroop task that involved naming the color of target words that were either related to the sentence context (the appropriate condition), related to the object noun in the sentence, but unrelated to the context (the inappropriate condition), or unrelated to any aspect of the sentence (the neutral condition). They found that the less-skilled readers experienced greater interference relative to the neutral condition with targets that were both appropriate and inappropriate, while skilled readers experienced greater interference relative to the neutral condition only with appropriate targets. Merrill et al. (1981) argued that their evidence suggested that less-skilled readers encoded word meanings while reading a sentence that were both semantically related and unrelated to the

sentence context, while the skilled readers encoded only contextually relevant meanings.

A possible extension of the Merrill et al. (1981) findings, is that both skilled and less-skilled readers could be induced with instructions to process the words in sentences either as semantically integrated units or as individual word units. For example, in order to induce integrative processing, subjects could be instructed to construct an image that represents the meaning of each sentence. With imagery instructions, both skilled and less-skilled readers would be expected to show more color-naming interference when the context of the sentence is related than when the sentence context is unrelated. A different set of instruction could be used to induce readers to process words in a sentence as individual units. Subjects could be instructed to search through the sentence for a word that rhymes with a previously identified word. In this instructional condition both the skilled and less-skilled subjects would be expected to show more color-naming interference when the context of the sentence is both appropriate and inappropriate to the target word than when the sentence context is neutral.

An important assumption underlying both the Merrill et al. (1981) study and the proposed study involving instructions is that less-skilled readers have comprehension problems that go beyond decoding ability.

Evidence will be reviewed suggesting that less-skilled readers demonstrate at least two types of problems that are related to their failure to process sentences as integrated units. The first of these problems concerns organizing text into meaningful units and seems to be specific to processing written text. The second problem concerns the construction of semantic representations of both written and aurally presented text.

This introduction is organized in five major sections that describe the background research and provide a justification for the present research. First, an examination of the evidence that text organization is a print-specific problem for some readers is provided. In the second section literature is reviewed that demonstrates that less-skilled readers can also have problems constructing semantic representations of text. In the third section evidence is presented that successful decoding is necessary but not sufficient for successful comprehension. The role of context is discussed in this section and the Merrill et al. (1981) experiment is described in detail.

A discussion of why inducing readers to process text as integrated units might facilitate text comprehension is presented in the fourth section. This section also examines different methods used to induce readers to process words as integrated units or as individual units.

An overview of the study, along with the specific hypotheses and predictions, concludes the introduction. This fifth section also includes a review of the Stroop effect and several studies that have used Stroop-like tasks.

Evidence for text organization problems

There is a body of research that seems to show that some less-skilled readers are capable of decoding individual words, but are unsuccessful at comprehending text because they process words as individual units instead of organizing text into larger meaning units (e.g., Cromer, 1970, Fleisher, Jenkins, & Pany 1979; Kendall & Hood, 1979; Levin, 1973; Martinez, Ghatala, & Bell, 1980; Oaken, Wiener, & Cromer, 1971; Steiner, Wiener, & Cromer, 1971). This research seems to have originated with a series of studies conducted by Cromer and his colleagues (Cromer, 1970; Oaken, Wiener, & Cromer, 1971; Steiner, Wiener, & Cromer, 1971). Two of these studies and related research will be reviewed.

Cromer (1970) identified two groups of college students with reading deficiencies. The first group demonstrated inadequate vocabulary skills and was identified as a Deficit group. The other group demonstrated a word-by-word organizational strategy for processing text and was called the Difference group. Both of these groups were matched with skilled readers. The

skilled readers were matched with the Deficit group on IQ scores and with the Difference group on both IQ and vocabulary test scores. Comprehension performances were compared in four modes of text presentation: word-by-word, normal sentence, meaningful segments, and non-meaningful segments. The major finding was that the Difference group performed as well as the skilled readers in the meaningful segment condition, but significantly less well in the other three conditions. A comparison of the Deficit groups' performance across presentation conditions revealed that they performed best in the word-by-word condition although their performance was still below that of the skilled readers.

Oaken, Wiener, & Cromer (1971) compared comprehension performances of skilled and less-skilled fifth-grade readers in auditory and visual conditions that were either well organized or poorly organized. They found that the listening comprehension performance of skilled readers did not vary as a function of organization in the auditory condition, but that reading comprehension performance was lower in the poorly-organized visual condition than in the well-organized visual condition. The less-skilled readers performed poorly when the auditory presentation was poorly organized, but they performed as well as the skilled readers when the presentation was well organized. For the visual condition, Oaken et al. (1971) first established a

base-line level of reading comprehension performance. Oaken et al. then trained the less-skilled readers to identify the words that would appear in subsequent passages. Despite the identification training, the less-skilled readers performed more poorly than the skilled readers on the reading comprehension test. The performance of the less-skilled readers seemed unaffected by the identification training. The general conclusion from these findings was that the less-skilled reader's comprehension performance suffered because of a failure to impose organization on either the poorly organized auditory stimuli or the visually presented text.

A similar conclusion concerning less-skilled readers and visually presented text was drawn by Fleisher, Jenkins, & Pany (1979). In their study, less-skilled readers were trained to rapidly decode words both in isolation and in phrases. They found that comprehension performance did not improve following the rapid decoding training. They also found that the less-skilled readers decoded words in context at the same rate as they decoded single words, while the skilled readers decoded words in context more rapidly than they decoded single words. Fleisher et al. argued that this finding suggested that the less-skilled readers were processing the text in a word-by-word manner.

One general conclusion concerning the less successful comprehenders who do not organize text has been that they fail to utilize the syntactic cues inherent in sentences. Fletcher (1981) reviewed evidence that suggests that by the fourth grade, readers commonly utilize syntactic cues to organize groups of words into meaningful units. A number of researchers have noted that less-skilled readers in the elementary grades demonstrate an insensitivity to grammatical constraints when processing written text (Clay & Imlach, 1971; Fletcher, 1981; Gibson & Levin, 1975; Isackson & Miller, 1976; Kendall & Hood, 1979; Weinstein & Rabinovitch, 1971).

Problems related to constructing semantic representations

While it may be that text organization problems tend to be specific to written text, other researchers have identified more general language comprehension problems related to the ability to construct a memory representation of text that is sufficiently integrated and specific to the text. There is evidence that these problems can occur at several different levels of text processing. For example, researchers have found that representation problems can occur at the proposition and sentence levels (e.g., Oakhill, 1983; Tierney, Bridge, & Cera, 1978-1979; Townsend, Carrithers, & Bever, in press) and at the intersentence and more global thematic levels (e.g., Garnham, Oakhill, & Johnson-Laird, 1982; Perfetti &

Goldman, 1976; Smiley, Oakley, Worthen, Campione, & Brown, 1977; Tierney, Bridge, & Cera, 1978-1979). This evidence will be reviewed below.

Oakhill (1983) compared the performances of skilled and less-skilled seven to eight year old readers on a cued sentence recall task. The sentences were presented aurally. The cues were either original nouns from the sentence or instantiated nouns. The original nouns were superordinate category names (e.g, fish, tool, furniture), while the instantiated nouns were more specific category members (e.g., shark, saw, chair). The sentences were constructed such that the instantiated cues were implied by the context. For example, the cue for the sentence The tool cut through the wood was saw. Oakhill found that the less-skilled readers performed more poorly than the skilled readers on the recall task only when the cue was an instantiated noun. This finding suggests that the skilled readers more often than the less-skilled readers utilized context in order to construct an integrated and specific semantic representation of aurally presented sentences.

Tierney, Bridge, and Cera (1978-1979) examined the story recall performances of skilled and less-skilled third grade students. Their subjects read two stories aloud and then recalled the first story. They found that the less-skilled readers recalled fewer propositions and

the propositions they did recall were less complete. Furthermore, when they examined whether subjects recalled propositions within a logical, they found evidence that the less-skilled readers did not appear to process the logical connections between propositions. These findings suggest that some less-skilled readers have problems constructing and connecting propositions.

Townsend, Carrithers, & Bever (in press) compared performances of skilled and average readers from the sixth to eighth grade and college levels on reading and listening tasks that involved constructing and connecting propositions. A general finding was that the average readers performed more poorly than the skilled readers on both reading and listening tasks. For example, the average readers performed more poorly than the skilled readers on comprehension tasks involving listening to single sentences, listening to complete stories, and reading stories. Townsend et al. used a Meaning Probe task to examine a listener's ability to access the meaning of a sentence and a Word Probe task to examine how a listener processes structural relations between clauses. They additionally had a Find-the-Odd-Word task that assessed syntactic processing and two tasks that examined thematic processing.

The evidence from Townsend et al. study suggested that the average school age readers processed sentences in

a word-by-word manner without integrating the words into a propositional structure. They found that the skilled school age readers constructed propositions, but they were less likely than college readers to connect propositions. The two groups of college readers were found to differ in terms of how they related propositions. The average college readers were found to connect propositions only through using schematic knowledge of the text, while the skilled college readers used structural, morphemic, and schematic information.

While Townsend et al. (in press) found differences between average and skilled school age readers in their ability to construct and connect propositions, Smiley, Oakley, Worthen, Campione, and Brown (1977) found differences between skilled and less-skilled seventh grade readers at a more global level of text processing. Smiley et al. found that the less-skilled seventh grade readers were less sensitive than skilled readers to the thematic relevance of different story elements. Subjects in the Smiley et al. study read one story and listened to another story. For both reading and listening presentations, the skilled readers recalled significantly more text elements that were highly important to the story. The recall performances of the less-skilled readers did not demonstrate awareness of different levels of importance in either mode of presentation. This finding seems to

suggest that these less-skilled readers did not construct representations of the stories that discriminated between highly important and less important story information. It seems possible that these less-skilled readers were struggling with lower level meaning analysis (i.e., propositional analysis) and that may have hindered analysis at a more global level.

Perfetti and Goldman (1976) examined third and fifth grade skilled and less-skilled readers' performances on a verbal memory probe task. The subjects listened to stories that were interrupted by a probe word. The probe word was either from a near or far main or subordinate clause in a preceding sentence. The subjects task was to state the word that came after the probe word in the story. The less-skilled subjects were, in general, less likely to recall the words. This finding seem to demonstrate that the less-skilled readers were not holding in working memory the text elements necessary for connecting propositions within and across sentences. Furthermore, it was found that when the probe was from a far clause the skilled readers would more often than the less-skilled readers recall an appropriate paraphrase. Perfetti and Goldman argued that this finding suggests that the less-skilled readers were not efficiently encoding interpretations of clauses.

Garnham, Oakhill, and Johnson-Laird (1982) also provided evidence that less-skilled readers were less likely to connect propositions across sentences. They examined recall performances of skilled and less-skilled seven to eight year old readers on three types of stories. The stories were either normal, randomly ordered, or randomly ordered with referential continuity restored and referents easily identifiable. They found that skilled readers would make the bridging inferences that were necessary for establishing the coherence of a story that was randomly ordered but had referential continuity intact. The less-skilled readers did not make the necessary connections.

Decoding ability: Necessary but not sufficient

The importance of decoding ability. The assumption that effective decoding skills are a prerequisite to successful reading comprehension is based on the notion that comprehension processes operate on the products of lexical access and that lexical access is primarily a stimulus-driven process (e.g., Gough, 1983; Onifer & Swinney, 1981; Seidenberg, Tannenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979). In other words, lexical access occurs as a result of decoding and independently of context effects. Swinney (1979) and Onifer and Swinney (1981) provided convincing evidence that lexical access during sentence processing was independent of effects of

semantic context. Their evidence was based on lexical decision tasks involving cross-modal priming. In both studies it was found that lexical decisions for ambiguous words both related and unrelated to a sentence context were facilitated when the decisions immediately followed the presentation of the ambiguity in the sentence. When the decisions were delayed, only the word related to the context was facilitated.

Onifer and Swinney (1981) additionally found that both the frequent and less frequent meanings of ambiguous words were facilitated immediately following their presentation in a sentence even when the sentence context was biased toward either of the meanings. These findings suggest that selection of word meanings occurs as a result of a post-access decision process. These experiments seem to suggest that the decoding process provides access to the lexicon and that currently activated semantic context guides the post-access decision process whereby an appropriate meaning is selected.

Evidence consistent with this view was provided by Seidenberg et al. (1982) in a series of five experiments designed to examine the processing of lexical ambiguities in sentences. Their evidence supported the notion that information concerning words (i.e., meanings, phonology, and orthography) is automatically accessed from the lexicon without influence from contextual information.

The role of context in comprehension. The findings on the autonomy of lexical access seem to indicate that decoding is the necessary first step to successful reading comprehension, but they also seem to suggest that effective decoding skills do not provide a sufficient condition for successful comprehension. If post-access meaning selection processes are guided by the semantic context that the reader has represented in memory, then the reader needs to have constructed a representation that is meaningfully organized and sufficiently integrated in order for the selection processes to operate effectively. The assumption here is that comprehension involves an ongoing process whereby a memory representation of the stimuli is constructed and continually updated as new information is processed. The semantic context that guides post-access selection processes is the part of that meaningfully integrated memory representation that is specific to the text.

It seems to follow from the above argument that a competent decoder who either fails to organize text into meaningful units or who fails to construct an integrated meaning representation may access the lexicon, but not benefit from post-access selection processes. If these selection processes cannot function, then the reader may

not encode word meanings that are specific to the text. Evidence for this phenomenon was provided by Merrill et al. (1981).

Merrill et al. (1981) examined the semantic interference effects demonstrated by good and poor comprehenders with a modified Stroop task. Fifth grade subjects read a sentence aloud and then were presented with a target word typed in one of four colors. The task was to name the color as quickly as possible. The sentence context was either appropriate, inappropriate, or neutral to a target. For example, the sentence The man played the piano was appropriate for the target word music and inappropriate for the target word heavy. The sentence The man moved the piano was appropriate for the target heavy and inappropriate for the target music. The sentence The girl felt the rain was neutral for the target word music and the sentence The girl heard the rain was neutral to the target heavy. The idea was that the amount of interference produced with sentences that were appropriate and inappropriate for the targets, relative to the amount produced with sentences that were neutral for the targets, would reflect the extent to which the meaning represented from the sentence included target meanings. Merrill et al. (1981) hypothesized that the good comprehenders would experience more interference relative to the neutral condition in only the appropriate sentence condition and

that poor comprehenders would experience more interference in both the appropriate and inappropriate sentence conditions. This is exactly what they found. They concluded that the poor comprehenders semantically encoded individual word meanings separately rather than as integrated units while reading sentences.

It should be noted that Merrill et al. (1981) also compared the color-naming latencies of skilled and less-skilled readers across sentence context and single word context conditions. In the single word condition only the object from each sentence was presented. They found the same pattern of results for both skilled and less-skilled readers. The words that were objects from either appropriate or inappropriate sentences were associated with more interference than the objects from neutral sentences. This finding supports the idea that the differences found in color-naming latencies between the skilled and less-skilled readers in the sentence context condition were due to differential use of the contextual information from the sentence.

Evidence that without contextual information skilled readers will represent distinctive properties of words was provided by Cairns, Cowart, and Jablon (1981) with college students using a probe latency task. The task followed either a predictive or nonpredictive sentence in which the target appeared. Cairns et al. (1981) argued that the

contextual information provided by a predictive sentence is used by post-access processes controlled by a Message Processor that integrates information from context, other relevant knowledge, and inferences in order to develop a 'conceptual representation' of the linguistic message. In their experiment, Cairns et al. (1981, experiment 3) found shorter latencies with the nonpredictive sentences. The interpretation was that since contextual information was not useful for representing the sentence target word in an integrative way, the salient features of the target were represented, thereby facilitating recognition of the target in the subsequent task.

It is interesting that the Cairns et al. (1981) interpretation of how a target word is encoded when the context is nonpredictive is similar to the interpretation provided by Merrill et al. (1981) concerning how the poor comprehenders encoded words while reading sentences. Merrill et al. argued that the poor comprehenders seemed to semantically encode words in a sentence as if they were isolated units. It seems that the poor comprehenders in the Merrill et al. study encoded words in a manner similar to the presumably competent readers in the Cairns et al. study when the context did not facilitate integration of word meanings. In short, the less-skilled readers in the Merrill et al. study failed to use contextual knowledge

to guide the selection and integration of word meanings that were appropriate to the sentence context.

Evidence that instructions can affect semantic encoding

The findings from the Merrill et al. (1981) study seem to suggest a reason why instructing readers to process words in sentences as integrated units may facilitate comprehension for competent decoders with comprehension problems. It is possible that instructions to integrate word meanings in a sentence induce readers to encode the words in an integrated representation. This, in turn, might facilitate the post-access selection processes. In other words, such instructions might induce the reader to encode the contextual information necessary to integrate word meanings and select text appropriate meanings.

Integrative procedures for processing words and text. It seems to follow, then, that there may be a number of procedures for inducing readers to process words in text as related units. Arguments for several procedures will be described. Gibson and Levin (1975) suggested that pictures that meaningfully depict units of related ideas in conjunction with text would demonstrate to the reader that the same type of organization is inherent in the text. They also described work by Frase (Gibson & Levin, 1975) that encouraged young readers to use organizational strategies. One study involved instructing children to

attend to specific, important concepts while reading and another had readers answer questions with the text present that required them to combine information across sentences.

A study conducted by Weaver (1979) was similar to the studies by Frase (Gibson and Levin, 1975) in that Weaver also instructed subjects in how to utilize organizational strategies. Weaver trained third grade students to understand how a sentence is organized with direct instruction and by having them group words to form sentences. They practiced their organizational skills with a sentence anagram task that included sentences with varying numbers of words. Weaver found that subjects who receive training outperformed the control group on a sentence anagram test, a cloze test and a sentence recall test.

Imagery has also been proposed as a strategy that facilitates comprehension. For example, Levin (1973) found that fourth grade poor readers who were competent decoders benefitted from imagery instructions. The instructions were to create a visual image of each sentence in a story. The criterion task was a test involving questions concerning the content and sequence of events in the story. Levin found that the poor readers performed as well as good readers in the condition with imagery instructions.

Further evidence in support of imagery instructions was provided by Ledger and Ryan (1985). They found that kindergarten subjects recalled pictograph sentences much better following training on integrative imagery. They concluded that even young children can learn to use imagery as an strategy that facilitates semantic integration.

Inducing both integration and non-integration of words. The studies described above demonstrate that there are a number of possible methods for inducing or teaching organizational strategies and strategies for constructing integrative semantic representations. There are also studies that have shown that readers can be induced with instructions to process words either as related or as individual units. Several studies of this type will be reviewed.

Martinez, Ghatala, & Bell (1980) had seventh grade good and poor readers perform an orienting task while reading. The task involved judging the pleasantness of words, sentences and paragraphs. The idea was that the orienting task would induce the readers to encode the chunk of text they were to judge. The findings of interest were that performance on a story recall task was better (compared to a control) for poor readers when they were judging sentences, while good readers in the sentence condition performed about the same as good readers in the

control condition. Both groups of readers performed more poorly than their controls in the word condition. The condition that encouraged integration at the sentence level was found to be optimal for story recall performance.

Seamon (1972) conducted an experiment that involved different instructions expected to induce subjects to process words either as separate units or as an integrated unit. Seamon examined response latencies on a short-term recognition memory task for words varying in set size. Subjects received one of three sets of instructions: 1) to subvocally rehearse each word, 2) to create and hold separate images of the individual words, or 3) to create an integrated image with all the words of a set. The first two types of instructions resulted in increased latencies for larger sets, while the integrative imagery instructions resulted in a constant latency across different set sizes. These differences in latencies seem to reflect differences in how the words within the sets were encoded as a function the type of instructions. Individual units were encoded for each word when subjects were instructed to attend to the salient features of each word, but when the subjects were instructed to create an image that involved all the words in a set a single unit was encoded that integrated the words meaningfully.

Overview of the study

Before the issue of instructional intervention can be addressed using the Merrill et al. (1981) paradigm, the paradigm must be subjected to a replication. The present study was designed for that purpose. The findings should suggest that skilled, school-age readers process the words in sentences as integrated units while less-skilled, school-age readers process the words as individual units.

Several modifications of the Merrill et al. study were made. One important difference concerns the stimuli seen by each subject. The Merrill et al. design involved repeated measures on both the variables context (full sentence and single word) and relatedness (inappropriate, appropriate, and neutral). Their subjects saw the same object nouns and targets at each level of both context and relatedness.

Merrill et al. generated six pairs of sentences each of which was either inappropriate or appropriate depending on which target they were paired with (e.g., The man moved the piano was appropriate for heavy and inappropriate for music). For the word context condition, the object noun from each sentence was paired with the targets in the same manner. Their subjects saw a related object (which was either appropriate or inappropriate for a given target) paired with a target eight times. The advantage to this procedure is that it requires a relatively small set of

stimuli. The problem is that there might be effects of viewing an object noun paired with a target that carry over to another trial where that same noun and target are paired. In other words, the outcome of viewing a noun paired with a target on a previous trial may affect the outcome of viewing that same pairing on a subsequent trial.

The possibility for such confounding effects was controlled for in the present study in two ways. First, the context variable was treated as a between-subjects variable. Secondly, subjects in both the single word and full sentence conditions encountered an object noun and a target only once. A considerably larger set of stimuli, patterned after the Merrill et al. stimuli, were developed to facilitate this control. (See the Materials sub-section for details concerning how the stimuli were developed.)

Fifth grade subjects were used since they were used by Merrill et al. (1981) and because there is evidence that by this grade level skilled readers process sentences as integrated structures. Fletcher (1981), for example, provided evidence that by the fourth grade the average reader commonly utilizes grammatical structure while reading. Futhermore, Paris and Lindauer (1976) provided evidence that by the fifth grade readers will construct representations of sentences that include aspects that

were implied by the context. This suggests that comprehension strategies are probably employed by skilled fifth grade readers without any instructional inducements.

The same modified-Stroop task as used by Merrill et al. was used in the present study. A brief review of the Stroop effect and how the original task has been extended should elucidate why a Stroop-like task would be appropriate for the present research question. Stroop (1935) found that when a subject had to name the color print of a word spelling the name of another color there was a reliable interference effect in that subjects took longer to name the color than subjects whose task was to name the color when it appeared in squares (i.e., without a word). This finding suggested that subjects automatically read the words with the result that the lexical entry for the word name was primed along with the entry for the color name. The Stroop-interference effect seems to reflect competition for selection between two primed responses.

This response competition effect is not limited to naming colors when color names are printed in conflicting colors. West and Stanovich (1978) noted that the same effect has been found when subjects have to name the color print of a word that was heard several seconds before. In general, whenever a response that conflicts with the color-naming task is primed visually or auditorially this

interference effect should be found because there will be two responses competing for selection.

Several studies have used modified versions of the Stroop task to examine whether context affects lexical access (e.g., Conrad, 1974; Oden and Spira, 1983; West and Stanovich, 1978). The general hypothesis for these studies was that if context facilitates lexical access, then the Stroop effect should be strongest (i.e, there should be greater color-naming interference) when the colored word is specific to the context. West and Stanovich (1978) found that this hypothesis was supported for fourth and sixth grade readers, but not for college students. Conrad (1974) found only a slight increase in interference for context specific words.

Oden and Spira (1983) argued that Conrad's findings suggested that degree of activation may be affected by context and that a related target would remain more strongly activated. They tested this hypothesis by delaying the Stroop-like task by 500 milliseconds (msec).[1] They found 100 msec more interference for the targets that were related to the context than for the targets that were unrelated to the context. They argued that their finding suggested that lexical access or initial activation of

---

[1]Conrad (1974) wrote that the color-naming task in her study immediately followed the sentence, but she did not explain the apparatus used. It seems possible that there was an unspecified delay.

lexical items may operate independently of context, but that context seems to affect a decision process that selects a context appropriate meaning.

For the present study, the plan was to also delay the Stroop-like task long enough so that the competing response should reflect context effects on post-access decisions regarding accessed words. The idea is that with a delay the lexical entries that are semantically related to the context encoded by the subject (either an integrated structure or individual word units) should compete for response selection with the color name. Since the purpose was to replicate the Merrill et al. (1981) findings, the same procedures for the Stroop-like task were followed. In that study the Stroop-like task was presented after a one second delay.

The following predictions concerning the sentence context condition are based on the findings of Merrill et al. (1981). In the neutral and appropriate conditions of sentence relatedness the skilled and less-skilled readers are not expected to perform differently on the color-naming task. It is predicted that the skilled and less-skilled readers will perform differently when the sentence relatedness is inappropriate. More specifically, the skilled readers are expected to show more interference compared to the neutral conditions only with the targets that are paired with appropriate sentences. The less-

skilled readers are expected to show more interference compared to the neutral conditions with targets that are paired with both appropriate and inappropriate sentences.

In order to be confident that the differences in color-naming latencies between the skilled and less-skilled subjects are due to differential use of context, the effects of a sentence context on the color-naming task were compared to the effects of a single word context on the same task. The expectation was that the single word condition findings would also replicate those of Merrill et al. (1981). The two groups of readers were not expected to perform differently in this condition. Both groups were expected to demonstrate approximately the same amount of interference with the appropriate and inappropriate words, and the interference was expected to be greater than what they demonstrated with the neutral words.

A naming task was included as a measure of decoding competency. In order for the Stroop paradigm to differentiate between readers who comprehend sentences and readers who do not comprehend the sentences, both the skilled and less-skilled readers must demonstrate some level of decoding proficiency. Furthermore, Pace and Golinkoff (1976) have demonstrated, with a similar semantic interference task, that subjects who are unable to decode the target words will not experience semantic

interference. Therefore, as a check on decoding ability, a naming task involving all the target words plus additional nonexperimental words was included in the experimental session. Measures of both decoding accuracy and vocalization latency were obtained. Differences between skilled and less-skilled readers were expected on the vocalization measure with the less-skilled readers expected to have longer latencies. Both groups of readers were expected to know most, if not all, of the words, but, as Perfetti has argued (1985), vocalization latency will often distinguish skilled from less-skilled readers when accuracy measures do not.

CHAPTER 2

METHOD

Design

The experiment involved a mixed design with two between-subjects variables with two levels each and one within-subjects variable with four levels. The between-subjects variables were reading ability (skilled and less skilled) and context (full sentence and single word). The within-subjects variable was context/target relatedness (appropriate, inappropriate, neutral 1, and neutral 2). The experimental conditions are shown in Table 1. Each of the sentences and words was tested at each of the four levels of relatedness, but each subject viewed different sentences or words at the four levels of relatedness.

Table 1

Conditions and Levels

Reading Ability	Context	Relatedness of Context/Target			
	Word	Approp.	Inapp.	Neutral1	Neutral2
Skilled					
	Sentence	Approp.	Inapp.	Neutral1	Neutral2
	Word	Approp.	Inapp.	Neutral1	Neutral2
Less Skilled					
	Sentence	Approp.	Inapp.	Neutral1	Neutral2

Subjects

Subjects were 46 fifth-grade students from two elementary schools in a Western Massachusetts school district. The population of the district is predominantly white and middle class. Reading ability measures on the elementary school students from the district (e.g., placement in the basal series and subtest scores on standardized reading tests) are best characterized by a bimodal distribution with a considerably larger number of students in the higher modal group.

The school committee agreed to having fifth-grade students solicited for participation in the study through a parental permission process. Parental permission forms were given to the five fifth grade teachers who were instructed by their principals to distribute the forms to their students. Although 75 permission forms were distributed to the school with the larger 5th grade population, only 19 forms were returned from that school and 16 of those students actually participated in the experiment. At the second school, 30 out of 47 forms were returned with parental permission. All of those students participated in the study.[2]

---

[2]The experimenter did not meet with the teachers from the first school because they would not agree to a meeting. The experimenter learned later that the teachers were in fact resistant to the idea of having the study done in their school. It seems likely that not all of the 75 forms were actually distributed and it is more than likely that the teachers did not

The students were classified as skilled or less-skilled readers based on several pieces of information. Information concerning reading grade-level placement at the end of the fourth grade was available for the 19 students from the first school. Grade level placements were largely determined by a student's progress through the basal reading series. Teachers assigned a grade-level placement of 4.8 for those students who were reading at-or-above grade level at the end of the fourth grade. Teachers assigned grade-level placements of either 2.3, 3.8, or 4.3 for those students who were reading below-grade level at the end of the fourth grade. For the 30 subjects from the second school, grade-level placement for current instruction was provided by the teachers in the form of who was receiving instruction (i.e., from the basal series) at-or-above the fifth grade level and who was receiving instruction geared to below the fifth grade level.

Information was provided on students from both schools concerning who was receiving special reading services. Percentile rankings on the reading comprehension subtest of the Metropolitan Achievement Test (MAT) were available from the third grade for 37 subjects. The median percentile rank on the MAT for these subjects was 69.25. Scores on the Gates-MacGinte were available

---

encourage students to return the forms.

for four subjects who were receiving remedial reading instruction. There were five subjects for whom no standardized test information was available. For two of these subjects the information was not available because they were new to the school system. The other three subjects were exempt from testing because they were receiving special education services.

For the subjects who had both a grade-level placement and a percentile ranking on the MAT subtest, the criteria for a skilled reader classification were an at-or-above grade-level placement and a percentile ranking greater than the median of 69.25. There were seven subjects with the requisite grade-level placement who did not meet the MAT subtest criterion for skilled-reading classification. The data from these subjects were excluded from all analyses. One subject was classified as skilled in the absence of a MAT ranking. The general criteria for a less-skilled reader classification was a below grade-level placement and a percentile ranking below the median. Only one subject with a below grade-level placement had a percentile ranking above the median. The data from this subject were not used.

Of the three subjects who did not have a grade-level placement because they were receiving their primary reading instruction through a special education program, only one was unable to complete the task. This subject's

data were excluded from all analyses. The other two subjects were classified as less skilled. A fourth subject was receiving special education services for reading instruction. This subject had a grade-level reading placement of 2.3 and was also classified as less-skilled. The Gates-MacGinte scores for the four remedial reading subjects along with the grade level information indicated that these subjects were reading an average of one year below grade level. These subjects were also classified as less skilled.

An additional seven students were classified as less skilled. For five of these subjects the classification was based on the convergence of evidence from their grade-level placement of at least .5 years below grade level and their comprehension subtest scores on the MAT. The other two subjects were receiving at-grade-level reading instruction, but their oral reading performances were flagged as "less skilled" by the experimenter who made comments on the data collection sheets concerning each subject's experimental session. Both of these subjects stumbled and hesitated while reading the experimental materials that had been developed for the fourth-grade level. All of the other skilled readers read the materials fluently and many of the other less-skilled readers read with a greater degree of fluency than these two subjects. In both of these cases the comprehension

subtest scores from the MAT were congruent with the subjective impression of the experimenter. One subject scored at the 42nd percentile while the other scored at the 20th percentile.

A total of 16 subjects were classified as less-skilled readers. These subjects averaged .8 years below the expected 4.8 grade level (this average includes the two subjects with at-grade-level placements and excludes the two subjects with special education placements instead of a grade-level indicator). One of these subjects' data were thrown out due to technical problems during the session.

There were a total of 22 subjects classified as skilled readers. The data for four of the skilled readers were thrown out. In one case the subject consistently gasped or sighed into the microphone before naming the color, making his color-naming reaction times meaningless. In the other cases there were technical problems during the session.[3] Of the remaining 18 skilled reader subjects, one did not have a test score because he/she was new to the school system as of the fourth grade. This subject was included as a skilled reader because of her

---

[3]At the first school, the only place to set up the apparatus was below a fan. For four subjects the voice-key continually picked up the fan. The experimenter was eventually able to adjust the sensitivity of the voice-key.

current grade-level placement and on the recommendation of her teacher.

The median percentile score on the MAT comprehension subtest for the less-skilled subjects was 36.5 and for the skilled it was 87.5. The median for the less-skilled subjects does not reflect the lowest ability subjects who were either exempt from testing or who took the Gates-MacGinte test because of their remedial reading status.

Materials

Sentences that consisted of a subject, a verb, and an object were used in the full sentence condition. The sentences not taken from the Merrill et al. study (1981) were constructed with vocabulary that was considered to be familiar to most students reading at the fourth-grade level. Appropriate vocabulary was chosen by reference to both Fry's (1972) list of 600 Instant Words and Dale and O'Rourke's (1976) vocabulary inventory. Fry's (1972) list is based on several word-frequency studies and consists of high-frequency words used in the first through fourth grade levels. The inventory constructed by Dale and O'Rourke (1976) provides a percentage score based on testing students' (at grade levels 4, 6, 8, 10, 12, 13, and 16) familiarity with different meanings of many different words. A score of 65% or greater at the fourth grade-level was the criterion for acceptability of words.

Pairs of appropriate and inappropriate sentence-target combinations were constructed such that they were similar to the sentence-target combinations used in the Merrill et al. (1981) study. Merrill et al. had six pairs of appropriate and inappropriate sentences that were identical except for the verb. The emphasized semantic feature of the object noun changed as a function of the verb. Each sentence in a pair was appropriate for one target and inappropriate for another target. In other words, there was a target that was related to the overall meaning of the sentence (appropriate) and a target that was related to the object noun but not related to the overall sentence meaning (inappropriate). For example, the sentence The girl fought the cat was appropriate for claw and inappropriate for fur. The sentence The girl touched the cat was appropriate for fur and inappropriate for claw. Thus each sentence-target combination is matched with two targets. Ninety-six pairs of sentence-target combinations were constructed in this manner.

Each of the 192 sentence-target combinations, along with the 12 used by Merrill et al., was rated by a panel consisting of 9 graduate students and one faculty member. The panel rated each sentence-target combinations on the degree to which the sentence context was related to the target. A seven point scale was used with 1 indicating unrelated and 7 indicating extremely related. The

criterion for an appropriate sentence-target combination was a mean rating of 5 or greater. The criterion for an inappropriate sentence-target combination was a mean rating of less than 4. In order for an appropriate and inappropriate sentence-target combination pair to be used it had to meet the criteria for two targets (with the appropriate/inappropriate relationship reversed). There were 56 pairs that met the criteria. The remaining 40 pairs were discarded.

Merrill et al. also had six pairs of neutral sentences that were identical except for the verb. These neutral sentence pairs were unrelated to the appropriate and inappropriate pairs (i.e., there was a different subject, verb, and object). They had one neutral sentence for each pair. For the present study, there were two neutral sentences for each target. These neutral pairs had the same subject and object and the object was different from the object in the appropriate and inappropriate sentence pair. One of the neutral sentences contained the verb from the appropriate sentence and the other contained the verb from the inappropriate sentence. For example, the two neutral sentences for the sentences about the girl and the cat were <u>The girl fought the snake</u> and <u>The girl touched the snake</u>. Both of these sentences were used as neutral sentences for the targets <u>fur</u> and

claw. The were 56 neutral sentence pairs constructed in this manner.

For each target, then, there was a sentence quadruple corresponding to the four levels of relatedness. The same quadruple was paired with two targets. The sentences that were appropriate and inappropriate were reversed for the two targets as were the verbs in the two neutral sentences. Thus there were 28 sentence octaves since for each sentence there were two appropriates, two inappropriates, two neutral1s, and two neutral2s. The complete set of full sentence stimuli and their corresponding conditions of context-target relatedness is provided in Appendix A.

Each subject encountered two sentences from each octave. One sentence was paired with one target, the other sentence was paired with the other target. A subject encountered a different level of context-target relatedness with each target and a different noun object. See Appendix B for an example of how the assignment of four subjects in the sentence context condition to the conditions of relatedness was accomplished for the first two sentence octaves. An important difference between the present study and the Merrill et al. study is that subjects in the present study saw each of the 56 targets only once and 56 different noun objects.

For the word context condition, the object from each sentence was paired with the target. Thus there was a word pair for each target. Each word pair was then paired with two targets. For example, the context words cat and snake formed the pair for the two targets fur and claw. The appropriate versus inappropriate distinction was not actually meaningful in the single word condition in terms of context-target relatedness since the single words were either related or neutral to the targets. For example, cat was related to both the targets fur and claw, while snake was neutral to both the targets fur and claw. The appropriate versus inappropriate distinction was retained in order that the single word condition was comparable to the full sentence after the full sentence data was collapsed over the two verbs and targets into the four levels of relatedness.

The target words were presented in four different colors for the Stroop task. The colors were blue, green, red, and purple. Straight color-naming latencies were obtained for the colors (presented as a list in a random order) with six adult subjects. This was done to rule out the possibility that latencies when naming purple would be consistently longer. The means were 550 msec. for purple and 566 for the others. The assignment of colors to the targets was randomized on each trial.

The naming task included all the target words and additional, filler words. There were a total of 96 words. The filler words for the naming task were also chosen from Fry's lists and the Dale and O'Rourke's (1976) vocabulary inventory. The words were randomly arranged in a list format for the naming task.

Apparatus

The stimuli were presented on a Zenith color monitor with a portable Zenith 160 microcomputer. The MetraByte CTM-05 counter-timer and I/O expansion board was installed to accommodate a voice key and two response buttons. The voice key was used to measure vocalization latencies on the naming task and response latencies on the Stroop task. A microphone was connected to the voice key. The response buttons were used by the experimenter to record responses as correct or incorrect and to initiate the onset of trials.

Procedure

Subjects were randomly assigned to the context conditions except that an attempt was made to have an equal number of less-skilled readers in the single word and full sentence conditions and an equal number of skilled readers in the two context conditions. Random assignment was accomplished in the following manner. All subjects had been assigned a number and prior to data collection subject numbers were assigned to context

conditions. The subject numbers and context conditions were then recorded on data collection sheets. Teachers were given two cards stapled together. On the top card the subject's name was recorded and on the other card was the subject's number. Teachers were instructed to separate the cards and send students out of the classroom with the only the subject number card. The subject gave the card to the experimenter who matched it up with a data collection sheet.

When the subject arrived for the experiment he/she sat approximately .5 meters from the monitor with the center of the screen at eye level. The subject was first oriented to the equipment and told that there were two parts to the experiment. (The experimenter followed a written set of instructions for each condition and these instructions are included in Appendix C). The subject was told that the microphone would pick up extraneous noises and that he/she should try to sit quietly during both parts of the experiment. He/she was told that the study was concerned with whether or not children can read words and sentences quickly from a computer screen.

After the general instructions and orientation, the naming task was explained. The subject was instructed to read each word aloud as soon as it appeared on the screen and was told that both speed and accuracy would be measured. The speed measures were the vocalization

latencies recorded with the voice key that was interfaced with the computer. Accuracy measures were recorded by the experimenter who pressed the right button for a correct response and the left button for an incorrect response.

Following preliminary instructions, the subject completed 15 practice trials. Before beginning the 96 experimental naming trials, the subject was told that there would be a break half way through the trials to give them a chance to catch their breath. They were told to look for the message "Take a Break" on the screen and to sit quietly until they were ready to begin and then nod to the experimenter to indicate that they were ready.

After the subject completed the naming task, the Stroop task was explained. Depending on whether the subject had been assigned to the single word or full sentence condition, the subject was told that she/he would read either single words or sentences that would be followed by a single word presented in one of four colors. The subject was instructed to read the first word or sentence aloud and focus on the meaning and then name the color of the second word as soon as it appeared on the screen. (See Appendix C for the actual instructions.)

The subject was told that the computer would measure the speed of her/his color naming response and that the experimenter would record whether or not she accurately read the words or sentences and the colors. The

experimenter used the data collection sheet to record correct and incorrect responses on both the context stimuli questions and the color-naming task. The experimenter also recorded trials on which there were microphone problems (i.e., trials where either the microphone picked up extraneous noise or failed to pick up the subject's voice response).

The subject was also told that it was important to read the word or sentence preceding the color word carefully and for meaning. In order to encourage subjects to process the context stimuli, simple yes/no questions about the preceding word or sentence were inserted periodically after the color was named. For the single word condition, the questions were of the type Was it a ...?, and then a category would be named that would for half the questions be a correct category for the word. The same type of question was used in the sentence condition along with questions of the type Did the person . . .something? A verb would be named that for half of the questions would have been present in the sentence. It was important to vary the focus of the questions for the sentence condition so that subjects would need to attend to the entire sentence in order to reliably answer the questions correctly. The experimenter did not provide any information about how often the questions would occur. The questions occurred every fifth trial for the first 28

trials and every fourth trial for the second 28 trials. The frequency was varied in an attempt to keep the subject from learning the pattern and, in fact, no subject seemed to learn the pattern.

The subject was then shown the four colors on the computer screen and told that they were blue, green, red, and purple. The experimenter asked whether the subject agreed with those color names. None of the subjects expressed a problem with the color names identified by the experimenter. The subject was told that it was very important that they use the same color names through out the experimental session.

The subject then completed 24 practice trials before the experimental trials. The 56 experimental trials were presented in two blocks of 28 with a break in between. Subjects were allowed to decide the length of the break. The majority of subjects were ready to continue after about a minute. The entire experimental session took an average of 30 minutes.

CHAPTER 3

RESULTS AND DISCUSSION

Word-Naming Task

Effects of ability differences on word-naming latencies. A one-way analysis of variance (ANOVA) was employed to look at the effect of ability differences on word-naming latencies. A significant effect was found, $F(1,31) = 18.63$, $p. < .0002$. The mean word-naming latency for the less-skilled readers was 823 msec. and for the skilled readers it was 554 msec.

Effects of ability differences on proportion correct on word-naming task. Differences between the two ability groups on proportion correct on the naming task were also examined with a one-way ANOVA. The mean proportion correct was .92 for the less-skilled readers and .99 for the skilled readers. Since these data were extremely skewed, the one-way ANOVA was computed for the arc sine of the proportion correct scores. A significant difference between the two ability groups was found, $F(1,31) = 21.07$, $p. < .001$.

Context Stimuli Questions

A two-way ANOVA was computed to examine differences between skilled and less-skilled readers in the two context conditions (single word and full sentence) on the arc sine of proportion correct scores on the 12 context stimuli questions. The arc sine transformation was again

used because the data were very skewed. The mean proportion correct was .95 for the skilled readers and .84 for the less-skilled readers. Even though the purpose of the 12 context questions was to encourage all subjects to attend to the context stimuli and to demonstrate that they had in fact processed the content, a marginally significant main effect for ability was found, $F(1,29) = 4.07$, $p = .05$. There was no effect due to the different contexts and no interaction.

Oral Reading of Context Stimuli

An effect for ability was also found when a two-way ANOVA was used to examine the effects of ability and context differences on the arc sine of proportion correct scores on oral reading of the context stimuli. Since these data were also very skewed, the mean was .99 for the skilled readers and .95 for the less-skilled readers, the arc sine transformation was again used. The difference between the two ability groups was significant, $F(1,29) = 14.757$, $p < .001$. Again, there was no evidence for a context effect or interaction.

Color-Naming Latencies and Error Rates

Effects of two verbs in full sentence context condition. While the single word context condition had four levels of context-target relatedness (appropriate, inappropriate, neutral1, and neutral2), the full sentence context condition had eight levels since there were two

verbs that appeared in the appropriate, inappropriate and two neutral conditions of context-target relatedness. In order to compare the full sentence condition with the single word condition, it was necessary to first rule out any effects due to the two different verbs on both color-naming latencies and error rates in the full sentence condition. It should be noted here that only correct-trial latencies were included in any of the mean latencies. Since differences were expected for the different levels of relatedness, separate contrasts were carried out between the two verb conditions for each of the four levels of relatedness. None of the contrasts were significant for the latency data. The eight levels of context-target relatedness were then collapsed into four (appropriate, inappropriate, neutral1, and neutral2) in order to compare the color-naming latencies in the full sentence condition with color-naming latencies in the single word condition.

The error rate data were consistently very skewed in that both ability groups had relatively few errors in color-naming. For example, the overall mean error rate for the less-skilled readers was 8.5% while the overall mean for the skilled readers was 5.5%. Therefore, the appropriate ANOVA tests were computed with the error rate data transformed into the arc sine of the proportion of

errors. For the purpose of interpretation, though, the observed mean error rates will be reported.

The same contrasts were done with the error rate data as were done with the latency data in order to identify effects due to the two verb. A significant interaction with ability was found for the first neutral condition contrast, $F(1,16) = 5.04$, $p. < .05$. With the first verb, both ability groups had mean error rates of 6%. With the second verb, though, the skilled readers had a mean of zero while the less-skilled had a mean of 12%. So the skilled readers had no errors at all with the second verb, but the less-skilled readers had twice as many errors with the second verb than with the first verb. Since none of the marginal means were significantly different, the eight levels of context-target relatedness collapsed into four (appropriate, inappropriate, neutral1, and neutral2) in order to compare error rates in the full sentence condition with those in the single word condition.

Effects of two neutrals. The next set of analyses examined the effects of the two different neutral conditions on both color-naming latencies and error rates. Contrasts between the two neutrals were carried out in order to test for differences between the two neutrals and interactions with ability and context. For the latency data, there was no main effect and there were no interactions with either ability or context. The mean

latency for neutral1 was 935 and for neutral2 it was 939. The two neutrals were then collapsed into a single level of neutral relatedness.

For the error rate data, a significant interaction with context was found, $F(1,29) = 4.46$, $p. < .05$. For both ability groups in the single word condition there were higher error rates in the second neutral condition. The mean error rate for the less-skilled readers in the neutral1 condition of the single word condition was 3.5%, while in neutral2 the mean was 10%. The mean error rate for the skilled readers in the neutral1 condition of the single word condition was 4%, while in neutral2 the mean was 7%. Since the marginal means for the two neutrals were not statistically significant (the means were 5% for neutral1 and 6% for neutral2), the two neutrals were collapsed into a single level of neutral relatedness.

Effects of ability, context, and relatedness. The data were analyzed using a 2 (ability) X 2 (context) X 3 (relatedness) analysis of variance with repeated measures on relatedness. Looking first at the latency data, there were significant main effects for both ability and context. The mean for the less-skilled readers was 1070 msec. and the mean for the skilled readers was 843 msec., $F(1,29) = 27.79$, $p < .0001$. Regarding the effect for context, both skilled and less-skilled readers were faster when the preceding context was a full sentence rather than

a single word. The mean for the single word condition was 1003 msec. and for the full sentence condition it was 910 msec., $F(1,29) = 4.6$, $p. < .05$.

The means and standard deviations for each of the 12 cells are shown in Table 2 (see page 51). Contrasts between the appropriate and neutral mean, inappropriate and neutral mean, and appropriate and inappropriate mean were carried out for each level of ability and context, except for the skilled reader level of the full sentence condition where there clearly were no differences. None of the observed differences were statistically significant.

A different pattern of results was found for the error rate data. The means and standard deviations for the error rate data are shown in Table 3 (see page 52). A main effect for relatedness was found. The marginal mean for the appropriate condition was 9%, the marginal mean for the inappropriate condition was 6%, and the marginal mean for the neutral condition was 6%, $F(2,58) = 3.74$, $p. < .03$. The interaction between ability and relatedness was also significant, $F(2,58) = 3.10$, $p. < .05$. The less-skilled readers in both the sentence and single word context conditions had significantly more color-naming errors in the appropriate.

TABLE 2

Means and Standard Deviations for Color-Naming Latencies as a Function of Ability, Context, and Relatedness.

	SENTENCE CONTEXT		WORD CONTEXT	
	LESS SKILLED	SKILLED	LESS SKILLED	SKILLED
APPROP	1037	796	1125	863
	(99)	(115)	(208)	(124)
INAPP	1018	803	1148	912
	(73)	(115)	(176)	(138)
NEUTRAL	1007	800	1086	883
	(105)	(115)	(138)	(145)

NOTE. Latencies are in milliseconds. Standard deviations appear in parentheses.

TABLE 3

Means and Standard Deviations for Error Rates as a Function of Ability, Context, and Relatedness.

	SENTENCE CONTEXT		WORD CONTEXT	
	LESS SKILLED	SKILLED	LESS SKILLED	SKILLED
APPROP	13	4	13	7
	(9)	(5)	(12)	(6)
INAPP	7	7	4	5
	(7)	(9)	(7)	(7)
NEUTRAL	7	4	7	6
	(8)	(6)	(7)	(5)

NOTE. Error rates are percentages. Standard deviations appear in parentheses.

Item Analyses

All of the above analyses were repeated with the items functioning as the random variable instead of the subjects. For this set of analyses the context factor is the only between-items factor as the ability factor is a within-items factor. Thus there were 56 items per case and 56 cases. The findings with items will be reported in the same sequence as the subjects' results. Since the rationale for each analysis is the same as is was for the subjects' analyses, the rationales will be omitted. The findings that are discrepant with the subjects' analyses will be noted.

Effects of two verbs in full sentence context condition. Separate contrasts were carried out between the two verb conditions for each of the four levels of relatedness (appropriate, inappropriate, neutral1, and neutral2) on both the latency and error rate data. None of the contrasts were significant for either the latency or error rate data. The interaction with ability for the first neutral condition with the error rate data found in the subjects' analysis approached but did not reach significance in the items' analysis, $F(1,27) = 3.88$, $p. < .06$. The eight levels of context-target relatedness were collapsed into the four (appropriate, inappropriate, neutral1, and neutral2) in order to compare the full sentence condition with the single word condition.

Effects of two neutrals. The effect of the two different neutral conditions on both color-naming latencies and error rates were next analyzed with contrasts between the two neutrals. For the latency data, there was no main effect and there were no interactions with either ability or context. The mean latency for neutral1 was 955 msec. and for neutral2 it was 948 msec.

For the error rate data, there was no main effect, but a significant interaction with context was found, $F(1,54) = 5.87$, $p. < .02$. Both the skilled and less-skilled readers in the single word condition had higher error rates with neutral2. The mean error rate for the less-skilled readers in the neutral1 condition of the single word condition was 3.6%, while in neutral2 the mean was 10%. The mean error rate for the skilled readers in the neutral1 condition of the single word condition was 4%, while in neutral2 the mean was 7%. The two neutral conditions were collapsed into a single level of neutral relatedness for both the latency and error data.

Effects of ability, context, and relatedness. The data were analyzed using a 2 (context) X 2 (ability) X 3 (relatedness) analysis of variance with repeated measures on ability and relatedness. Main effects for both context and ability were found with the latency data. The mean for the full sentence condition was 930 msec. and for the single word condition it was 997 msec. The difference was

significant, $F(1,54) = 126.7$, $p. < .001$. The mean for the less-skilled readers was 1062 msec. and for the skilled readers it was 865 msec. This difference was significant, $F(1,54) = 132.47$, $p. < .0001$.

A main effect for context was also found with the error data. This effect was not found with the subjects' data. The mean error rate was 4% for the full sentence condition and the mean error rate for the single word condition was 7%. This difference proved to be significant, $F(1,54) = 7.47$, $p. < .01$. The main effect for relatedness found with subjects was not found with items. The interaction between relatedness and ability found with the subjects' data was also significant in the item analysis, $F(1,54) = 5.64$, $p. < .005$. The marginal mean for less-skilled ability group in the appropriate condition was 10%, while the marginal mean for the skilled ability group in the appropriate condition was 5.5%.

Conclusions

The results clearly do not replicate the Merrill et al. (1981) findings. In the present experiment the skilled and less-skilled readers differed on the Stroop task only in terms of overall response latencies. The more perplexing finding, though, was the absence of an effect due to context-target relatedness. Since this finding is anomalous with the previous research involving Stroop tasks, it was decided that a follow-up study was

indicated to determined why the expected context-target relatedness effect was absent in the present experiment.

The follow-up study

The purpose of the follow-up study was to explore whether the color-naming latencies of 26 college students participating in the same modified Stroop-task as used in the present study would follow the pattern indicated by previous findings with modified-Stroop tasks (e.g., Conrad, 1974; Merrill et al., 1981; Oden and Spira, 1983; West and Stanovich, 1978) or the pattern found with the fifth grade students in the present study. Evidence for a pattern consistent with previous research would be found if the adult subjects had longer latencies with targets that were related to the context stimulus relative to their latencies with targets that are either unrelated (i.e., inappropriate in the full sentence condition) or neutral to the context stimulus. Evidence for a pattern consistent with the current findings would be found if the color-naming latencies did not vary as a function of context-target relatedness. Reading ability was not a factor in the follow-up study since the purpose was to look at the more general effects of context-target relatedness.

Effects of context and relatedness with adult data. There were no effects found for either the two verbs or two neutral so the data were collapsed into a 2 (context)

X 3 (relatedness) analysis of variance with repeated measures on relatedness. For both the latency and error rate data, there were no main effects of either context or relatedness and no interaction. For the latency data, though, the interaction between context and relatedness approached significance, $F(2, 48) = 2.74$, $p < .08$.

The means and standard deviations for the latency data are shown in Table 4. It can be seen from Table 4 that in the single word condition there is a difference of approximately 32 msec between the related levels of relatedness (i.e., the appropriate and inappropriate levels) and the neutral level. A contrast on these data revealed a significant difference, $F(1,24) = 6.55$, $p < .02$. In other words, subjects in the full sentence condition did not demonstrate differences in color-naming latencies as a function of context-target relatedness while the subjects in the single word condition had longer latencies with related targets relative to the neutral targets.

TABLE 4

Means and Standard Deviations for Adult Color-Naming Latencies as a Function of Ability, Context, and Relatedness.

	SENTENCE	WORD
APPROP	665	706
	(137)	(148)
INAPP	671	708
	(122)	(125)
NEUTRAL	676	675
	(102)	(94)

NOTE. Latencies are in milliseconds. Standard deviation appear in parentheses.

CHAPTER 4

GENERAL DISCUSSION

The two features of the results that are most important are (a) the absence of differences in color-naming latencies as a function of context-target relatedness for either fifth grade ability group and (b) the failure to replicate the Merrill et al. (1981) findings concerning differences between skilled and less-skilled readers. The results concerning error rates and color-naming latencies that are directly relevant to the Merrill et al. findings will be discussed after the preliminary findings have been reviewed. The more general problem concerning the absence of an effect of context-target relatedness will be discussed after the comparisons between Merrill et al. and the present study have been made. This chapter will conclude with a discussion of the follow-up study and the methodological factors that may have influenced the present findings.

Word-Naming Task

The skilled and less-skilled readers were expected to differ on the naming task only in terms of vocalization latency since only words that were familiar to most fourth grade students were used. In fact, though, significant differences were found on both the latency and accuracy measures. The less-skilled readers were not only slower to name the words they were also less likely to name the

words correctly. It should be noted, though, that their mean performance was actually quite high (92% correct) and it seems safe to say that the words were generally familiar to them and fairly easy for them to name. The naming task findings also seem to show that the ability differences between the two groups of readers were substantial.

Context Stimuli Questions

Differences between the two ability groups were also found on the context stimuli questions. However, there was no effect of context and no ability x context interaction. The context questions were seemingly simple YES or NO questions concerning the sentence or word the subject had just read prior to naming the color of the target word. These questions were included as an inducement for subjects in both context conditions to process the context stimuli at the semantic level and to demonstrate that subjects had actually processed the content of the context stimuli. The skilled readers had a mean proportion correct score that exceeded the less-skilled readers' scores by 11%, but both groups performed with a relatively high degree of accuracy (95% and 84%, respectively).

This result is especially important in light of the fact that neither ability group demonstrated longer latencies with related targets relative to the neutral

targets. Had performance on these questions been relatively poor, then it could be argued that the subjects had not processed the context stimuli at the semantic level. Such a finding with the context questions would have provided a possible explanation for the flat effect across the levels of context-target relatedness. Given that subjects did perform well on the context questions, that explanation cannot be used to explain the absence of different latencies at the different levels of relatedness.

Merrill et al. (1981) included a recognition task involving simple line drawings that depicted the content of the context stimuli either accurately or inaccurately. They found substantially fewer errors on their task than found in the present study. Only 7 of their subjects made one error each. There were more errors in the present study probably because subjects had to evaluate whether the object noun was a category member for all the single word questions and half of the full sentence questions. To make such an evaluation probably requires more cognitive ability than does the task of verifying whether or not a picture means the same as what was depicted by a word or in a sentence.

Oral Reading of Context Stimuli

Proportion correct scores on oral reading of the context stimuli were obtained in order to assess whether

or not the stimuli were appropriate in terms of vocabulary level for both ability groups. While a significant main effect for ability was found with these proportion correct scores, both groups performed with a very high degree of accuracy. The means were 95% correct for the less-skilled and 99% correct for the skilled readers. It seems fair to conclude that the stimuli were in fact accessible to both groups of readers.

Error Rates on Color-Naming Task

In the Merrill et al. (1981) study, there were no consistent trends in the error rate data. In the present study, though, there was an interaction between relatedness and ability. An examination of Table 3 reveals that the less-skilled readers in both the single word and full sentence conditions had considerably more errors with appropriate targets while the skilled readers demonstrated little variation in error rates across the levels of relatedness.

The interpretation of this finding is not altogether straightforward. For subjects in the full sentence condition it could be argued that the increased difficulty with the appropriate target was due to the fact that the targets were related to the sentence context while the inappropriate and neutral targets were not related. This argument cannot be made, though, for subjects in the single word condition since both appropriate and

inappropriate targets were related to the context words. In other words, there is no reason for subjects in the single word condition to have more difficulty with appropriate targets relative to inappropriate targets.

The interaction was also found when items were treated as the random variable. For the lower ability group, the higher error rates were associated with the appropriate targets. In addition, a main effect for context was present in the item analyses in that there were more errors in the single word condition.

Color-Naming Latencies

Several unexpected results were observed when the effects of ability, context, and the three levels of context-target relatedness were examined. First, the less-skilled readers had significantly longer latencies across the levels of both context and relatedness. Secondly, there was a main effect for context with the longer latencies found in the single word condition. Thirdly, there was no effect of the variable context-target relatedness on the color-naming latencies. All three of these findings are discrepant with the Merrill et al. (1981) results. Each of result will be discussed in turn.

Merrill et al. (1981) did not find a main effect for ability and one was not predicted for the present study. It is interesting to note that the mean latency for the

less-skilled readers in the present study was over 100 msec longer than the mean for the less-skilled readers in the Merrill et al. study. Furthermore, the skilled readers in the present study were faster than the skilled readers in the Merrill et al. study by over 50 msec.

The differences between the subjects in the two studies cannot be easily explained in terms of different methods of classifying subjects as skilled or less-skilled readers. In both cases several indices of ability were use. Merrill et al. relied on different reading subtest scores from the Stanford Achievement Test. In the present study comprehension subtest scores on the MAT were used in conjunction with teacher grade-level assignments made on the basis of progress through the basal reading series. In both studies there was a clear distinction between the two groups in that the less-skilled readers had clearly performed below grade level on the indices used and the skilled readers had performed at-or-above grade level. It should also be noted that the number of subjects used in each study was also similar. Merrill et al. (1981) had 14 skilled and 14 less-skilled readers and in the present study there were 18 skilled and 15 less-skilled readers.

The difference in overall color-naming latency found between skilled and less-skilled readers in the present study might have been explainable had indices of performance in areas other than reading been obtained.

Several of the less-skilled readers were receiving special education services for reading and four were receiving remedial reading services (the difference between the two types of services has to do with the perceived severity of the disability with special education services provided for the more severe disability). It is possible that some of these subjects were receiving services for deficits in other domains as well. It is also possible that some of the less-skilled readers who were not receiving reading-related services were receiving services for other problems. The point is that information concerning ability in other domains or concerning general cognitive ability was not obtained whereas such information might help explain the difference in color-naming times found between the skilled and less-skilled readers.

As in the present study, Merrill et al. (1981) also found an unexpected main effect of context, but in their study the longer latencies were found in the full sentence condition. They argued for the possibility that, when the context was a sentence, greater processing capacity was required to hold the context stimulus in memory until the end of the trial. As a result, there would be less capacity left for processing the target which might result in longer latencies. One problem with this explanation is that it seems to imply that the one second delay was not a sufficient interval for sentence comprehension to occur.

If this implication were to be taken seriously, then the interpretation of the other findings concerning the full sentence condition cannot easily be made in terms of sentence comprehension.

The purpose here is not so much to cast doubt on the interpretation of the Merrill et al. findings as it is to demonstrate that the context effect found in their study did not lend itself to an obvious interpretation. The interpretation of why longer latencies occurred in the present study following a single word stimulus rather than following a full sentence is also problematic. A possible explanation concerns the salient similarity between the single word stimulus and the target stimulus that is not present in the full sentence condition.

It seems possible that subjects were distracted or confused by the similarity between the context and target stimuli when both were single words. When the word appeared in the normal white lettering as a context stimulus their task was to read it aloud, but when the word appeared in one of the four other colors they now had to name the color. Certainly there is a greater possibility for confusion between the context and target stimulus in the single word condition. It may be that there was a greater need for conscious allocation of attention in order to respond with a color name one second after reading aloud the word than there was when naming

the color after having read aloud a more distinct context stimulus such as a sentence.

The absence of an effect of context-target relatedness

The problem concerning the interpretation of the absence of any effect due to the context-target relatedness is far more serious and troublesome than the previous explanations concerning the unexpected main effects of ability and context. This is because an important assumption underlying the present study was that Stroop task latencies should vary depending on whether a target was related or unrelated to the context stimulus for all subjects who are capable of comprehending the context stimulus. This assumption was based on the studies described earlier by Conrad (1974), Oden and Spira (1983) and West and Stanovich (1978). So, regardless of whether or not the differences between skilled and less-skilled readers found by Merrill et al. could be replicated, it was assumed that, at the very least, the skilled readers would show longer latencies with the related-to context targets relative to the neutral targets. In the present study there were no differences in latencies for either group of fifth grade readers with related targets relative to the unrelated (i.e., inappropriate targets with full sentences) and neutral targets.

The modified Stroop tasks that were described in the introduction were somewhat different from the task used by Merrill et al. and in the present study in that those studies were addressing the issue of context effects on lexical access. It is unlikely, though, that the differences in the purposes of the tasks can help explain the absence of effects due to context-target relatedness found in the present study. Furthermore, a study conducted by Whitney, McKay, Kellas, and Emerson (1985) with college students showed an effect of context-target relatedness. That study involved a task that was very similar to the Merrill et al. paradigm for the full sentence stimuli except that the sentences were presented aurally.

Whitney et al. (1985) varied both the amount of delay (0, 300 msec and 600 msec) and the frequency (low, high) of the property of the noun object that was emphasized by the sentence context in addition to context-target relatedness (appropriate, inappropriate, and neutral). The relevant comparison to the Merrill et al. (1981) and present studies concerns the 600 msec delay condition. They replicated the pattern found by Merrill et al. with skilled readers only with the low-frequency stimuli. In other words, in the low frequency condition subjects had longer latencies with appropriate targets relative to the neutral targets while the latencies with inappropriate

targets were not longer. Interestingly, Whitney et al. found that in the high frequency condition subjects had longer latencies relative to the neutral targets with both the appropriate and inappropriate targets. This finding suggests that high-dominant properties of object nouns remain active for competent readers regardless of the extent to which the sentence context primes that property.

While the Whitney et al. (1985) study provides new insights concerning the effects of context-target relatedness, it also provides further evidence that an effect of context-target relatedness should have been found in the present study. The evidence indicated that with a delay of 600 msec or more longer latencies should be observed with competent readers with targets that are related to the context relative to the latencies observed with neutral targets. It should be noted that in the Whitney et al. study the stimuli were different for each trial as they were in the present study.

The follow-up study

In light of all the evidence indicating that an effect of context-target relatedness should be found, a follow-up study involving college students and the same methodology seemed necessary to help clarify the issue of why an effect of context-target relatedness was not found in the present study. There seemed to be at least three factors that may have resulted in the flat effect across

the levels of relatedness. One such factor was that, despite the attempt to include subjects who were very similar to those used by Merrill et al., the flat effect might be peculiar to the subjects used in the present study. A second possible factor was that the stimuli used in the present study may not have captured the desired manipulation of context-target relatedness. A third factor was that there might be aspects of the present methodology other than the stimuli that affected the current results.

It was hoped that the follow-up study would differentiate between a possible explanation due to subjects and a possible explanation due to methodology. Support for the explanation that the finding was peculiar to the subjects used in the present study would be present if the adult color-naming latencies followed the pattern expected based on the previous Stroop-task research. If, on the other hand, the pattern found with adult subjects replicates the pattern found with fifth grade subjects, then this would be support for an explanation based on the present methodology.

Unfortunately, the results with adult subjects did not provide unequivocal evidence for an explanation for the absence of any effect of context-target relatedness found in the present study with the fifth grade subjects. Instead, the results of the follow-up study were somewhat

consistent with the previous Stroop-task research and somewhat consistent with the current results. More specifically, the results from the single word condition were consistent with the previous research and the results from the full sentence condition replicate those found in the present study.

It does seem, though, that the results of the follow-up study do not strongly support an explanation based on differences between the fifth grade subjects in Merrill et al. and the present study. If there were something exceptional about the fifth grade students used in the present study, it is unlikely that the data from the adult subjects would look as similar as it does to the data from the fifth grade subjects. In light of the similarities, it seems more prudent to look to methodological factors for an explanation.

The possibility that the stimuli are indicated in the absence of an effect of context-target relatedness is difficult to reconcile with the fact that the stimuli used were based on those used by Merrill et al. (1981) and were in fact very similar to those used by Whitney et al. (1985) except that the vocabulary level was lower for the present study. Furthermore, the appropriate and inappropriate stimuli used in the present study were rated in terms of their relatedness to the targets before being selected for use. While the neutral stimuli were not

subjected to a formal rating process, they seemed obviously unrelated to the target by adults reviewers. Nonetheless, in order to know for sure that the stimuli are not at the root of the problem, it would be necessary to compare the effects of different stimuli (i.e., the stimuli used in this study and stimuli used in one of the studies that found the effect) within the context of an experiment.

Another possible explanation is that the resolution of the color monitor may not have been as high as is necessary for a Stroop task. It is important to note that none of the reviewed studies involving modified-Stroop tasks used a computer monitor to present the stimuli. Instead, most of them utilized slide projectors (Conrad, 1974; Merrill et al., 1981; Oden and Spira, 1983; West and Stanovich, 1978) and Whitney et al. used a tachistoscope to present the Stroop stimuli.

It seems possible that a monitor resolution problem could manifest itself in a Stroop task in such a way that the color would be available before the word would come into focus. Some support for this effect could be the finding that the skilled readers in the present study were approximately 55 msec faster on average than the skilled readers in the Merrill et al. study. A more dramatic difference was observed between the adults in the present study who had a mean latency of 683 msec and the adult

subjects in the 600 msec delay condition of the Whitney et al. study who had a mean latency of 951 msec. In other words, the color-naming latencies found with the competent readers in the present experiments seemed to be considerably shorter than what has been observed in other Stroop-task experiments.

It seems clear that the Stroop task paradigm used in the present studies should not be used for similar purposes until the problems encountered here have been clarified and resolved. One possible next step to addressing this problem would be to design an experiment that would compare the Stroop effects found when the task is presented on a color monitor like the one used in the present studies with the effects found with the more often used slide projector apparatus. If the results found with the computer-presented task replicated those found in the present study and the effects found with the slide projector replicate the typical findings, then this would be evidence that a standard color monitor is not the best method for displaying a Stroop task, at least with the current state of the technology. It seems likely that any deleterious effects on experiments involving computer-presented stimuli will soon be overcome by the availability of high-resolution monitors. In the meantime, researchers should be careful not to assume that computer-presentation of stimuli will not introduce additional error variability into the data.

APPENDICES

Appendix A

Full Sentence Condition Stimuli

Context-Target Relatedness	Sentence	Target
Appropriate	The girl touched the cat.	fur
Inappropriate	The girl fought the cat.	fur
Neutral	The girl touched the snake.	fur
Neutral	The girl fought the snake.	fur
Neutral	The girl fought the snake.	claw
Neutral	The girl touched the snake.	claw
Appropriate	The girl fought the cat.	claw
Inappropriate	The girl touched the cat.	claw
Inappropriate	The boy held his nose.	sniffle
Appropriate	The boy blew his nose.	sniffle
Neutral	The boy held his horn.	sniffle
Neutral	The boy blew his horn.	sniffle
Neutral	The boy blew his horn.	smell
Neutral	The boy held his horn.	smell
Inappropriate	The boy blew his nose.	smell
Appropriate	The boy held his nose.	smell
Appropriate	The woman used her broom.	floor
Inappropriate	The woman flew her broom.	floor
Neutral	The woman used her kite.	floor
Neutral	The woman flew her kite.	floor
Neutral	The woman flew her kite.	witch
Neutral	The woman used her kite.	witch
Appropriate	The woman flew her broom.	witch
Inappropriate	The woman used her broom.	witch
Appropriate	The boy sat near the fire.	warm
Inappropriate	The boy saw the fire.	warm
Neutral	The boy sat near the stone.	warm
Neutral	The boy saw the stone.	warm
Neutral	The boy saw the stone.	smoke
Neutral	The boy sat near the stone.	smoke
Appropriate	The boy saw the fire.	smoke
Inappropriate	The boy sat near the fire.	smoke

Inappropriate	The man moved the piano.	music
Appropriate	The man played the piano.	music
Neutral	The man moved the card.	music
Neutral	The man played the card.	music
Neutral	The man played the card.	heavy
Neutral	The man moved the card.	heavy
Inappropriate	The man played the piano.	heavy
Appropriate	The man moved the piano.	heavy
Appropriate	The boy watched the movie.	screen
Inappropriate	The boy ate during the movie.	screen
Neutral	The boy watched the game.	screen
Neutral	The boy ate during the game.	screen
Neutral	The boy ate during the game.	popcorn
Neutral	The boy watched the game.	popcorn
Appropriate	The boy ate during the movie.	popcorn
Inappropriate	The boy watched the movie.	popcorn
Inappropriate	The man fell in the snow.	slippery
Appropriate	The man drove on the snow.	slippery
Neutral	The man fell in the road.	slippery
Neutral	The man drove on the road.	slippery
Neutral	The man drove on the road.	cold
Neutral	The man fell in the road.	cold
Inappropriate	The man drove on the snow.	cold
Appropriate	The man fell in the snow.	cold
Appropriate	The boy heard the duck.	quack
Inappropriate	The boy saw the duck.	quack
Neutral	The boy heard the lion.	quack
Neutral	The boy saw the lion.	quack
Neutral	The boy saw the lion.	swim
Neutral	The boy heard the lion.	swim
Appropriate	The boy saw the duck.	swim
Inappropriate	The boy heard the duck.	swim

Inappropriate	The man needed his glasses.	break
Appropriate	The man dropped his glasses.	break
Neutral	The man needed his comb.	break
Neutral	The man dropped his comb.	break
Neutral	The man drooped his comb.	see
Neutral	The man needed his comb.	see
Inappropriate	The man drooped his glasses.	see
Appropriate	The man needed his glasses.	see
Inappropriate	The man used the phone.	ring
Appropriate	The man heard the phone.	ring
Neutral	The man used the story.	ring
Neutral	The man heard the story.	ring
Neutral	The man heard the story.	call
Neutral	The man used the story.	call
Inappropriate	The man heard the phone.	call
Appropriate	The man used the phone.	call
Inappropriate	The boy finished the picture.	camera
Appropriate	The boy took the picture.	camera
Neutral	The boy finished the candy.	camera
Neutral	The boy took the candy.	camera
Neutral	The boy took the candy.	painting
Neutral	The boy finished the candy.	painting
Inappropriate	The boy took the picture.	painting
Appropriate	The boy finished the picture.	painting
Appropriate	The girl cooked the pumpkin.	pie
Inappropriate	The girl carved the pumpkin.	pie
Neutral	The girl cooked the chicken.	pie
Neutral	The girl carved the chicken.	pie
Neutral	The girl carved the chicken.	face
Neutral	The girl cooked the chicken.	face
Appropriate	The girl carved the pumpkin.	face
Inappropriate	The girl cooked the pumpkin.	face
Appropriate	The man caught the fish.	hook
Inappropriate	The man cooked the fish.	hook
Neutral	The man caught the apple.	hook
Neutral	The man cooked the apple.	hook
Neutral	The man cooked the apple.	fry
Neutral	The man caught the apple.	fry
Appropriate	The man cooked the fish.	fry
Inappropriate	The man caught the fish.	fry

Appropriate	The boy kept the mouse.	cage
Inappropriate	The boy caught the mouse.	cage
Neutral	The boy kept the toad.	cage
Neutral	The boy caught the toad.	cage
Neutral	The boy caught the toad.	trap
Neutral	The boy kept the toad.	trap
Appropriate	The boy caught the mouse.	trap
Inappropriate	The boy kept the mouse.	trap
Appropriate	The woman took an airplane.	trip
Inappropriate	The woman flew an airplane.	trip
Neutral	The woman took a flag.	trip
Neutral	The woman flew a flag.	trip
Neutral	The woman flew a flag.	pilot
Neutral	The woman took a flag.	pilot
Appropriate	The woman flew an airplane.	pilot
Inappropriate	The woman took an airplane.	pilot
Appropriate	The boy watched the bird.	fly
Inappropriate	The boy heard the bird.	fly
Neutral	The boy watched the bus.	fly
Neutral	The boy heard the bus.	fly
Neutral	The boy heard the bus.	sing
Neutral	The boy watched the bus.	sing
Appropriate	The boy heard the bird.	sing
Inappropriate	The boy watched the bird.	sing
Appropriate	The man ate the corn.	dinner
Inappropriate	The man picked the corn.	dinner
Neutral	The man ate the banana.	dinner
Neutral	The man picked the banana.	dinner
Neutral	The man picked the banana.	field
Neutral	The man ate the banana.	field
Appropriate	The man picked the corn.	field
Inappropriate	The man ate the corn.	field

Appropriate	The woman planted flowers.	garden
Inappropriate	The woman painted flowers.	garden
Neutral	The woman planted bushes.	garden
Neutral	The woman painted bushes.	garden
Neutral	The woman painted bushes.	colors
Neutral	The woman planted bushes.	colors
Appropriate	The woman painted flowers.	colors
Inappropriate	The woman planted flowers.	colors
Appropriate	The woman returned the book.	library
Inappropriate	The woman finished the book.	library
Neutral	The woman returned the dress.	library
Neutral	The woman finished the dress.	library
Neutral	The woman finished the dress.	read
Neutral	The woman returned the dress.	read
Appropriate	The woman finished the book.	read
Inappropriate	The woman returned the book.	read
Appropriate	The girl enjoyed her school.	learn
Inappropriate	The girl saw her school.	learn
Neutral	The girl enjoyed her lunch.	learn
Neutral	The girl saw her lunch.	learn
Neutral	The girl saw her lunch.	building
Neutral	The girl enjoyed her lunch.	building
Appropriate	The girl saw her school.	building
Inappropriate	The girl enjoyed her school	building
Appropriate	The woman cleaned her teeth.	brush
Inappropriate	The woman used her teeth.	brush
Neutral	The woman cleaned her desk.	brush
Neutral	The woman used her desk.	brush
Neutral	The woman used her desk.	chew
Neutral	The woman cleaned her desk.	chew
Appropriate	The woman used her teeth.	chew
Inappropriate	The woman cleaned her teeth.	chew

Appropriate	The boy felt the sun.	hot
Inappropriate	The boy drew the sun.	hot
Neutral	The boy felt the rain.	hot
Neutral	The boy drew the rain.	hot
Neutral	The boy drew the rain.	round
Neutral	The boy felt the rain.	round
Appropriate	The boy drew the sun.	round
Inappropriate	The boy felt the sun.	round
Appropriate	The girl heard the bee.	buzz
Inappropriate	The girl felt the bee.	buzz
Neutral	The girl heard the water.	buzz
Neutral	The girl felt the water.	buzz
Neutral	The girl felt the water.	sting
Neutral	The girl heard the water.	sting
Appropriate	The girl felt the bee.	sting
Inappropriate	The girl heard the bee.	sting
Appropriate	The man wanted the pizza.	hungry
Inappropriate	The man burned the pizza.	hungry
Neutral	The man wanted the letter.	hungry
Neutral	The man burned the letter.	hungry
Neutral	The man burned the letter.	oven
Neutral	The man wanted the letter.	oven
Appropriate	The man burned the pizza.	oven
Inappropriate	The man wanted the pizza.	oven
Appropriate	The girl liked the milk.	drink
Inappropriate	The girl spilled the milk.	drink
Neutral	The girl liked the cereal.	drink
Neutral	The girl spilled the cereal.	drink
Neutral	The girl spilled the cereal.	wet
Neutral	The girl liked the cereal.	wet
Appropriate	The girl spilled the milk.	wet
Inappropriate	The girl liked the milk.	wet

Appropriate	The girl enjoyed the ice cream.	eat
Inappropriate	The girl touched the ice cream.	eat
Neutral	The girl enjoyed the chair.	eat
Neutral	The girl touched the chair.	eat
Neutral	The girl touched the chair.	sticky
Neutral	The girl enjoyed the chair.	sticky
Appropriate	The girl touched the ice cream.	sticky
Inappropriate	The girl enjoyed the ice cream.	sticky
Appropriate	The woman wore the watch.	wrist
Inappropriate	The woman needed the watch.	wrist
Neutral	The woman wore the hat.	wrist
Neutral	The woman needed the hat.	wrist
Neutral	The woman needed the hat.	time
Neutral	The woman wore the hat.	time
Appropriate	The woman needed the watch.	time
Inappropriate	The woman wore the watch.	time
Appropriate	The man stayed in the hospital.	sick
Inappropriate	The man worked in the store.	sick
Neutral	The man stayed in the store.	sick
Neutral	The man worked in the store.	sick
Neutral	The man worked in the store.	doctor
Neutral	The man stayed in the store.	doctor
Appropriate	The man worked in the hospital.	doctor
Inappropriate	The man stayed in the hospital.	doctor

Appendix B

Assignment of Subjects to Conditions

An example of how the assignment of four subjects to the conditions of relatedness was accomplished for the first two sentence octaves. The actual order in which sentences were presented was randomized for each subject.

Sub	Sentence Relatedness	Sentence	Target
S1	Appropriate	The girl touched the cat.	fur
S2	Inappropriate	The girl fought the cat.	fur
S3	Neutral	The girl touched the snake.	fur
S4	Neutral	The girl fought the snake.	fur
S1	Neutral	The girl fought the snake.	claw
S2	Neutral	The girl touched the snake.	claw
S3	Appropriate	The girl fought the cat.	claw
S4	Inappropriate	The girl touched the cat.	claw
S1	Inappropriate	The boy held his nose.	sniffle
S2	Appropriate	The boy blew his nose.	sniffle
S3	Neutral	The boy held his horn.	sniffle
S4	Neutral	The boy blew his horn.	sniffle
S1	Neutral	The boy blew his horn.	smell
S2	Neutral	The boy held his horn.	smell
S3	Inappropriate	The boy blew his nose.	smell
S4	Appropriate	The boy held his nose.	smell

Appendix C

Instructions for Subjects

Single word Condition

I. General-Begin after you show them the equipment.

A. I'm trying to find out whether children can read and understand words quickly when they are shown on a computer screen. The microphone is connected to the computer so that the time it takes you to say items will be recorded by the computer. I will record on paper whether or not you say the items correctly.

B. Once we have started to record items with the computer, we both need to sit quietly because the microphone will pick up other noise. We only want it to pick up your voice as you are saying the items.

C. The session will be broken up into different tasks. I will explain each task to you and you will complete some practice trials before we start the recorded trials.

II. Naming Task

A. Your first task will be to name--as quickly as you can-- words that will appear on the screen. A list of words will be presented one at a time. The words will appear alone in the center of the screen. You should try to name each word as soon as it appears.

B. It is important that you only make one response for each word. So if you make a mistake or think you make a mistake, don't try to correct it, just go on to the next word.

C. Now I will show you some words for practice naming. Remember that you want to be both fast and correct.

D. AFTER THE PRACTICE TRIALS-- Do you understand the task? Now we are ready to begin the recorded trials. Half way through the list of words there will be a break. It will say "Take a Break" on the screen. This break is to give you a chance to catch you breath. Please sit quietly during the break and then nod to me when you are ready to begin.

III. Stroop/Color Naming Task--AFTER THE NAMING TRIALS

A. The next task I want you to do is probably more fun than the naming task, but it is also a little more complicated. You will first read aloud a word that will appear in white lettering on the screen. When you finish reading the word I will hit a button and one second later a word will appear in one of four colors. Your task will be to name the color that the word appears in as quickly as you can. For the first word, the one that is presented in white, you want to read for meaning as you normally read words. But when the second word appears in color, you want to focus your attention on naming the color.

B. In order to make sure that you are reading the first words carefully, I will sometimes ask you questions about the words after you have named the color of the second word. You will know when it is time for a question because it will say on the screen "Time for a question!" You will always answer either yes or no to the questions. If you really don't know the answer you should answer no. The questions will be easy if you are reading the words carefully. Let me give you an example of a question. Let's say that the word was "BED" A YES question would be "Was is a piece of furniture?" and a NO question would be "Was it a plant?" We are going to go through some practice trials, but first I want to show you the four colors.

C. As I show you the four colors I'm going to ask you if you agree with the name I give to each color. I'm doing this because it is really important that you use the same name for these colors through out the session. SHOW COLORS OK? Remember to use those color names all the time.

D. AFTER SHOW COLORS--Now we are ready for the practice trials. Remember that you will first read aloud the first word, the one that will appear in white lettering, then you will name the color of the single word that follows the sentence as quickly as you can. START PRACTICE TRIALS

E. AFTER PRACTICE TRIALS--Do you think you have the hang of it? Very good. Now we can begin the recorded trials. I've broken it up into two halves to give you break in between. Remember that once we get going you can't stop to ask questions or fix mistakes you need to wait for the break. But don't worry because it only takes a few minutes for each half. So are you ready?

F. DURING BREAK--This is the break. You are doing very well. I just want to remind you that it is very important that you name the color as quickly as you can. OK? Ready for the second half?

G. AFTER THE EXPERIMENT, THANK THE SUBJECT AND TELL HIM/HER THAT SHE/HE DID A GREAT JOB.

Instructions for Subjects
Full Sentence Condition

I. General-Begin after you show them the equipment.

A.I'm trying to find out whether children can read and understand words quickly when they are shown on a computer screen. The microphone is connected to the computer so that the time it takes you to say items will be recorded by the computer. I will record on paper whether or not you say items correctly.

B.Once we have started to record items with the computer we both need to sit quietly because the microphone will pick up other noise. We only want it to pick up your voice as you are saying the items.

C.The session will be broken up into different tasks. I will explain each task to you and you will complete some practice trials before we start the recorded trials.

II. Naming Task

A.Your first task will be to name--as quickly as you can-- words that will appear on the screen. A list of words will be presented one at a time. The words will appear alone in the center of the screen. You should try to name each word as soon as it appears.

B.It is important that you only make one response for each word. So if you make a mistake or think you make a mistake, don't try to correct it, just go on to the next word.

C.Now I will show you some words for practice naming. Remember that you want to be both fast and correct.

D.<u>AFTER THE PRACTICE TRIALS--</u> Do you understand the task? Now we are ready to begin the recorded trials. Half way through the list of words there will be a break. It will say "Take a Break" on the screen. This break is to give you a chance to catch you breath. Please sit quietly during the break and then nod to me when you are ready to begin.

III. Stroop/Color Naming Task--<u>AFTER THE NAMING TRIALS</u>

A.The next task I want you to do is probably more fun than the naming task, but it is also a little more complicated. You will first read aloud a sentence that will appear in white lettering on the screen. When you finish reading the sentence I will hit a button and one second later a word will appear in one of four colors. Your task will be to name the color that the word appears in as quickly as you can. For the sentence that is presented first and in white, you want to read for meaning as you normally read sentences. But when the single word appears in color, you want to focus your attention on naming the color.

B.In order to make sure that you are reading the sentence carefully, I will sometimes ask you questions about the sentence after you have named the color of the single word. You will know when it is time for a question because it will say on the screen "Time for a question!" You will always answer either yes or no to the questions. If you really don't know the answer, you should answer no. The questions will be easy if you are reading the sentences carefully. Let me give you an example of a question. Let's say that the sentence was "THE MAN SAT ON THE BED" A YES question would be "Was it about a piece of furniture?" and a NO question would be "Was it about a plant?" Another type of YES question would be "Did the person sit on something? and a NO question would be "Did the person clean something? We are going to go through some practice trials, but first I want to show you the four colors.

C.As I show you the four colors I'm going to ask you if you agree with the name I give to each color. I'm doing this because it is really important that you use the same name for these colors through out the session. SHOW COLORS OK? Remember to use those color names all the time.

D.AFTER SHOW COLORS--Now we are ready for the practice trials. Remember that you will first read aloud the sentence that will appear in white lettering, then you will name the color of the single word that follows the sentence as quickly as you can. START PRACTICE TRIALS

E.AFTER PRACTICE TRIALS--Do you think you have the hang of it? Very good. Now we can begin the recorded trials. I've broken it up into two halves to give you break in between. Remember that once we get going you can't stop to ask questions or fix mistakes you need to wait for the break. But don't worry because it only takes a few minutes for each half. So are you ready?

F.DURING BREAK--This is the break. You are doing very well. I just want to remind you that it is very important that you name the color as quickly as you can. OK? Ready for the second half?

G.AFTER THE EXPERIMENT, THANK THE SUBJECT AND TELL HIM/HER THAT SHE/HE DID A GREAT JOB.

REFERENCES

Cairns, H. S., Cowart, W., & Jablon, A. D. (1981). Effects of prior context upon the integration of lexical information during sentence processing. Journal of Verbal Learning and Verbal Behavior, 20, 445-453.

Clay, M. M., & Imlach, R. H. (1971). Juncture, pitch, and stress as reading behavior variables. Journal of Verbal Learning and Verbal Behavior, 10, 133-139.

Conrad, C. (1974). Context effects in sentence comprehension: A study of the subjective lexicon. Memory & Cognition, 2, 130-138.

Cromer, W. (1970). The difference model: A new explanation for some reading difficulties. Journal of Educational Psychology, 61, 471-483.

Dale, E., & O'Rourke, J. (1976). The living word vocabulary. Elgin, IL: Field Enterprises Educational.

Fleisher, L. S., Jenkins, J. R., & Pany, D. (1979). Effects on poor readers' comprehension of training in rapid decoding. Reading Research Quarterly, 15, 30-48.

Fletcher, J. M. (1981). Linguistic factors in reading acquisition: Evidence for developmental changes. In F. J. Pirozzolo & M. C. Wittrock (Eds.), Neurophysiological and cognitive processes in reading (pp. 261-294). New York: Academic Press.

Fry, E. B. (1972). Reading instruction for classroom and clinic. New York: McGraw-Hill.

Garnham, A., Oakhill, J., & Johnson-Laird, P. N. (1982). Referential continuity and the coherence of discourse. Cognition, 11, 29-46.

Gibson, E. J., & Levin, H. (1975). The psychology of reading. Cambridge, Massachusetts: MIT Press.

Gough, P. B. (1983). Context, form, & interaction. In K. Rayner (Ed.), Eye movements in reading (pp. 203-210). New York: Academic Press.

Isackson, R. C., & Miller, J. W. (1976). Sensitivity to syntactic and semantic cues in good and poor comprehenders. Journal of Educational Psychology, 68, 787-792.

Kendall, J. R., & Hood, J. (1979). Investigating the relationship between comprehension and word recognition: Oral reading analysis of children with comprehension disabilities. Journal of Reading Behavior, 11, 41-48.

Ledger, G. W., & Ryan, E. B. (1985). Semantic integration: Effects of imagery, enaction, and sentence repetition training on prereaders' recall for pictograph sentences. Journal of Experimental Child Psychology, 39, 531-545.

Levin, J. R. (1973). Inducing comprehension in poor readers: A test of a recent model. Journal of Educational Psychology, 65, 19-24.

Martinez, P., Ghatala, E. S., & Bell, J. A. (1980). Size of processing unit during reading and retention of prose by good and poor readers. Journal of Reading Behavior, 9, 89-95.

Merrill, E. C., Sperber, R. D., & McCauley, C. (1981). Differences in semantic encoding as a function of reading comprehension skill. Memory and Cognition, 9, 618-624.

Oaken, R., Wiener, M. and Cromer, W. (1971). Identification, organization, and reading comprehension for good and poor readers. Journal of Educational Psychology, 62, 71-78.

Oakhill, J. (1983). Instantiation in skilled and less skilled comprehenders. Quarterly Journal of Experimental Psychology, 35A, 441-450.

Oden, G. C., & Spira, J. L. (1983). Influence of context on the activation and selection of ambiguous word senses. Quarterly Journal of Experimental Psychology, 35A, 51-64.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. Memory & Cognition, 9, 225-236.

Pace, A. J., & Golinkoff, R. M. (1976). Relationship between word difficulty and access of single word meaning by skilled and less-skilled readers. Journal of Educational Psychology, 68, 760-767.

Paris, S. G. , & Lindauer, B. K. (1976). The role of Inference in children's comprehension and memory for sentences. Cognitive Psychology, 8, 217-227.

Perfetti, C. A. (1985). Reading Ability. New York: Oxford University Press.

Perfetti, C. A., & Goldman, S. R. (1976). Discourse memory and reading comprehension skill. Journal of Verbal Learning and Verbal Behavior, 14, 33-42.

Seamon, J. G. (1972). Imagery codes and human information retrieval. Journal of Experimental Psychology, 96(2), 468-470.

Seidenberg, M. S., Tannenhaus, M. J., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meaning of ambiguous words in context: Some limitations of knowledge-based processing. Cognitive Psychology, 14, 489-537.

Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J. C., & Brown, A. L. (1977). Recall of thematically relevant material by adolescent good and poor readers as a function of written versus oral presentation. Journal of Educational Psychology, 69, 381-387.

Steiner, R., Wiener, M., & Cromer, W. (1971). Comprehension training and identification for poor and good readers. Journal of Educational Psychology, 62, 506-513.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 18, 643-661.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. Journal of Verbal Learning and Verbal Behavior, 18, 645-659.

Tierney, R. J., Bridge, C. Cera, M. J. (1978-1979). The discourse processing operations of children. Reading Research Quarterly, 14, 539-573.

Townsend, D. J., Carrithers, C., Bever, T. G. (in press). Sentence processing in listening and reading among college- and school-age skilled and average readers. In R. Horowitz and S. J. Samuels (Eds.) Comprehending Oral and Written Language. New York: Academic Press.

Weaver, P. A. (1979). Improving reading comprehension: Effects of sentence organization structure. Reading Research Quarterly, 17, 131-146.

Weinstein, R., & Rabinovitch, M. S. (1971). Sentence structure and retention in good and poor readers. Journal of educational Psychology, 62, 25-30.

West, R. F., & Stanovich, K. E. (1978). Automatic contextual facilitation in readers of three ages. Child Development, 49, 717-727.

Whitney, P., McKay, T., Kellas, G., Emerson, W. A. (1985). Semantic activation of noun concepts in context. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11(1), 126-135.