

DOCUMENT RESUME

ED 252 813

CS 007 858

AUTHOR Omanson, Richard C.
TITLE Knowing Words and Understanding Texts: Models of Some Basic Relationships.
INSTITUTION Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
REPORT NO LRDC-1984/36
PUB DATE 84
NOTE 39p.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cognitive Processes; Elementary Education; Elementary School Students; *Language Processing; Memory; *Models; *Reading Comprehension; Reading Instruction; *Reading Processes; Reading Research; *Recall (Psychology); *Word Recognition
IDENTIFIERS *Text Processing

ABSTRACT

Prompted by current theories in reading suggesting that variation in word knowledge affects the processing of not only individual words but also clauses and sentences, this report explores some different ways in which children's comprehension processes may be affected by variation in word knowledge. The report first examines whether the effects of word knowledge on comprehension are similar for skilled and less-skilled readers, then compares how well stories containing unfamiliar versus recently taught words are remembered. It then considers various models of how comprehension processes may have been affected to produce the observed recall differences and evaluates them against experimental data. Finally, it presents some implications for theories of text processing, education, and the use of modelling as an experimental tool. (HOD)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED252813

LEARNING RESEARCH AND DEVELOPMENT CENTER

KNOWING WORDS AND UNDERSTANDING TEXTS:
MODELS OF SOME BASIC RELATIONSHIPS

1984/36

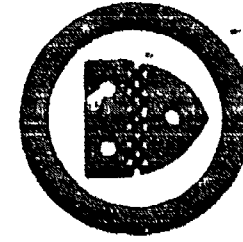
RICHARD C. OMANSON

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION

X

5007858

2



University of Pittsburgh

3

ERIC
Full Text Provided by ERIC

898 L0050

**KNOWING WORDS AND UNDERSTANDING TEXTS:
MODELS OF SOME BASIC RELATIONSHIPS**

Richard C. Omanson

Learning Research and Development Center
University of Pittsburgh

1984

The research reported herein was supported by the Learning Research and Development Center, funded in part by the National Institute of Education, U.S. Department of Education. The opinions expressed do not necessarily reflect the position or policy of NIE, and no official endorsement should be inferred.

It is not only whether a child knows the words in a text, but also how well the words are known, that affects comprehension.

knowing words and understanding texts: models of some basic relationships

richard c. omanson

Every day in elementary schools across the country, children read stories from basal reading lessons. For many children, reading these stories is not an easy task. When one considers the complexity involved in reading, this is not surprising. The ability to translate arbitrary squiggles of ink into an inferred representation of events and concepts involves a plethora of cognitive processes that must be simultaneously executed in a highly coordinated fashion. Reading involves coordinating perceptual processes that analyze letter shapes and letter combinations, phonological processes that bring to mind how the letters and the words they form sound, lexical processes that bring to mind the meanings of the words, and syntactic and semantic processes that analyze the meanings of the clauses and sentences that the words form.

Ideally, a school's reading program should be able to provide reading experiences that result in proficient reading skill for all children, including those who find reading difficult. In practice, schools fall short of this ideal. One reason schools are not able to make every child a proficient reader is that often children have trouble comprehending stories that are part of their reading lessons. If reading an assigned story is so difficult for children that they understand and retain

little from the experience, then it is likely that some of the processes involved in proficient comprehension are not sufficiently well-practiced to enable the reading process to proceed smoothly and efficiently. Reading under such circumstances is not likely to foster either the motivation to read or the kind of improvement in reading that will change the circumstances. A better situation would be for children to get a lot of practice at successfully understanding stories that gradually become more difficult and demanding as the abilities of the children improve. Thus, good instruction involves tailoring texts to the reading abilities of children, and successful tailoring requires knowing what makes reading difficult.

One factor that has a big influence on the difficulty of reading is the extent to which children are familiar with individual words. Most of the words encountered in basal reading selections are common words that are familiar to the children. However, basal reading selections also contain unfamiliar words that are included to give children practice at figuring out the meaning of words from context. They also contain words that have very recently been taught to the children. Current theories of reading suggest that such variation in word knowledge will affect comprehension because the component processes of reading interact with each other (cf. Just & Carpenter, 1980; Lesgold & Perfetti, 1981; Rumelhart, 1977). According to these theories, variation in word knowledge is thought to affect not only the processing of individual words, but also the processing of clauses and sentences that takes place during comprehension. What these theories have not spelled out, however, is exactly how comprehension processes change as a result of variation in word knowledge.

This chapter will explore some different ways in which children's comprehension processes may be affected by variation in word knowledge. It will also examine whether the effects of word knowledge on comprehension are similar for skilled and less-skilled readers. The way in which the effects of word knowledge on comprehension will be examined is to first compare how well stories containing unfamiliar versus recently taught words are remembered. Then, various models of how comprehension processes may have been affected to produce the observed recall differences

will be considered and evaluated against experimental data. Finally, some implications of the results for theories of text processing, education, and the use of modeling as an experimental tool will be considered.

effects of word knowledge on recall

The first question to address is whether comprehension differs for texts that contain unfamiliar versus recently instructed words. Comprehension refers to a person's understanding of what was stated in, and implied by, a text. There are a number of ways in which a person's understanding of a text can be assessed. Common procedures include answering questions (cf. Anderson & Biddle, 1975; Lehnert, 1978), composing summaries (e.g., Kintsch & van Dijk, 1978; Schank, 1975), and recalling the text from memory (e.g., Bower, 1978; Kintsch & van Dijk, 1978; Stein & Glenn, 1979). The method used to assess comprehension that will be considered here is recall. The rationale for using recall to assess comprehension is that recall reflects a person's representation of the text in memory that in turn is a result of the reader's particular understanding of the text (cf. Kintsch & van Dijk, 1978; Schank, 1975). Moreover, since the goal of reading is to gain information that can be used at a later time, and since recall is a conservative measure of the information that has been gained from a text, the use of recall as a measure of comprehension is particularly appropriate.

In order to assess the effects of unfamiliar versus recently taught words on recall, we will draw upon the results of a study reported by McKeown, Beck, Omanson, & Perfetti (1983), and its subsequent reanalysis by Omanson, Beck, McKeown, and Perfetti (1983). McKeown et al. taught inner-city fourth-grade children approximately 100 words through an unusually rich and intensive vocabulary program that extended over a five-month period. At the end of this instruction, the children were given two texts to read and recall from memory. One text, called the instructed-words story, was 270 words long and contained 30 of the instructed words. The other text, called the uninstructed-words story, was of similar length and structure and contained 30 unfamiliar words. By comparing the children's recall of the two stories, the effects of recently

4

Instructed versus unfamiliar words on comprehension could be compared. This within-subject comparison, however, is weakened by the fact that there are content and structural differences between the two stories in addition to the vocabulary differences. To ensure that any observed differences in the recall of the two stories were due primarily to the different types of vocabulary they contained, the two stories were given to a second group of fourth-graders who had not been taught any of the words. If the recall of both stories by the control classroom is similar to the recall of the uninstructed-words story by the experimental classroom, it is likely that the recall differences are due primarily to vocabulary differences. Thus, there was one situation in which children recalled a story containing instructed words (experimental children's recall of the instructed-words story), and three situations in which children recalled stories containing unfamiliar words (experimental children's recall of the uninstructed-words story, and the control children's recall of both stories).

In order to score the recalls of the stories, each story was divided into units of meaning called propositions (Kintsch, 1974), which will be described in more detail later. Each child's recall was then scored according to whether the gist of each story proposition appeared in the recall. In this way, the percentage of the propositions recalled by each child could be computed, enabling the performance of each group to be computed as the average percentage of propositions recalled by the children in the group.

The results of the scoring indicated that recall was highest for the story containing words that had been taught to the children. The children recalled about 20% of the story containing recently taught words but only recalled about 11% of the stories that contained unfamiliar words. Moreover, when the recalls of the children who were skilled in reading versus those who were less skilled were examined separately, the same pattern of results were obtained. For both skilled and less-skilled readers, about twice as much of the story containing recently taught words was recalled than the stories containing unfamiliar words.

The variation in word knowledge not only affected how much was recalled, but it also

affected which parts of the story were likely to be recalled. In examining which propositions were recalled, a characteristic pattern occurred. Propositions containing recently taught words were *more* likely to be recalled than propositions containing only familiar words. On the other hand, propositions containing unfamiliar words were *less* likely to be recalled than propositions containing only familiar words. In other words, the children were biased toward remembering the parts of the story containing the recently taught words, while, in contrast, were biased against remembering the parts of the story containing unfamiliar words.

In order to get a better sense of the qualitative differences between the recall of the story containing recently taught words and of the stories containing unfamiliar words, "prototypical" recalls were constructed to approximate the recall of the instructed-words story by the experimental group (who had been taught the instructed words) and by the control group (who were unfamiliar with the instructed words). The prototypical recalls were constructed by, first, rank ordering the propositions of the story according to the number of children in each condition that recalled them. Next, the mean recall for each group was rounded off to the nearest whole number of propositions, and then, that number of propositions were taken from the top of each group's rank ordering to create the prototypical recalls. The generated recalls are presented in Table 1.

Insert Table 1 about here

The plot of the instructed-words story centers on an ambitious violin novice, Sam, who gives a concert. A large woman interrupts Sam by talking to an acquaintance, and by eating food from a table being set for a party after the concert. Soon everyone begins to talk with each other and Sam walks off stage. Fortunately, a friend runs up to Sam and suggests that the audience play music with him. Instruments are found for the audience. Everyone plays and leaves feeling content.

As shown in Table 1, there were large differences between the prototypical recalls of this story by the experimental and control groups. The experimental group's prototypical recall of the instructed-words story contains twenty propositions, thirteen of which contained instructed words. It provides a good summary of the story in that it includes the initial setting that Sam, who was a novice at playing the violin, gave a concert, the conflict that he was repeatedly interrupted by a woman which caused him to walk away, and the resolution to the conflict offered by Sam's ally that the audience also played music which resulted in the audience being content (see Beck and Brewer's chapter in this volume for a discussion of discourse structure and what counts as a good summary of a story).

In contrast, the control group's prototypical recall of the instructed-words story contains twelve propositions, only one of which contains an instructed word. It does not provide as good a summary in that it omits from the setting the fact that Sam was a novice, it omits the entire conflict of Sam being interrupted while playing, and consequently, the audience also playing instruments is not depicted as a resolution to a conflict.

Given this demonstration that vocabulary instruction can affect text comprehension, it is important to note that the nature of the vocabulary instruction that is chosen may be important -- some types of vocabulary instruction may be much more successful than others in facilitating the comprehension of texts containing the instructed words. Indeed, of the few studies addressing the effects of vocabulary instruction on comprehension, some report improved comprehension (e.g., Beck, Perfetti, & McKeown, 1982; Draper & Moeller, 1971; Kameenui, Carnine, & Freschi, 1982; McKeown et al., 1983) while others report no difference (Jenkins, Pany, & Schreck, 1978; Tuinman & Brady, 1974).

In comparing the instruction used in these different studies, it appears that focusing on multiple aspects of vocabulary knowledge may be an important factor in whether vocabulary instruction affects comprehension (cf. Curtis & Glaser, 1983). For example, in the McKeown et al. study, the words were taught in sets of 8 to 10 words. Initially, the children received

instruction on each set of words for five days. On the first day, definitions were established and one or two activities presented that involved the children with the words. On the second day, the children wrote sentences for each word and engaged in some additional activity that provided fairly easy practice with word meanings. On the third day, the children were given activities in which they generated contexts containing new words. The purpose of these activities was to broaden the student's understanding of the words by promoting the establishment of relations between new words and already known words. The fourth day included an exercise that encouraged the children to think about the words in new ways and an activity in which the children repeatedly matched words and definitions while being timed. This latter activity was designed to increase fluent access to word meanings. On the fifth day, students took a multiple-choice test on the set of words studied. In addition, there were review sessions in which the words received practice activities. Thus, a typical set of words received about 40 encounters with the word in a highly varied instructional environment.

This type of instruction is in sharp contrast to that found in basal reading programs. The best situation in these programs occurs when a new vocabulary word is (1) introduced in a sentence that elucidates its meaning; (2) encountered in the reading selection which (3) prompts the child to look up its definition in the glossary; and (4) appears in an independently completed after-reading activity (Beck, McKeown, & McCaslin, 1983). Thus, in this best situation, the children only receive four, very limited encounters with each new word.

effects of word knowledge on comprehension processes

Since the recall data from McKeown et al. demonstrated that stories with unfamiliar versus recently taught words are comprehended differently, we can now address *how* the comprehension processes are affected. Describing comprehension processes is difficult. One can observe and describe differences in an outcome of comprehension such as recall. However, one cannot in the same way observe the comprehension processes directly to determine their nature or how they differ in different situations. What can be done is to construct different models of what the

processing might entail, and then examine which model best predicts the observed pattern of recall. This was the approach adopted by Omanson et al. (1983). The models that were constructed by Omanson et al. were based upon a model of comprehension processing presented by Kintsch and van Dijk (1978). In the next section, Kintsch and van Dijk's theory will be outlined. Then, within the framework of this theory, the results of analyses by Omanson et al. will be presented in which two models of comprehension of texts with unfamiliar words and two models of comprehension of texts with recently instructed words were evaluated.

kintsch and van dijk's model. Kintsch and van Dijk's model of comprehension is concerned with the comprehension processes that operate on the meaning of clauses and sentences. According to their view, comprehension begins by deriving a conceptual, or propositional, representation of each sentence in the text. A proposition consists of a predicate, or relation, and one or more arguments. Arguments can be either concepts within propositions or other propositions. For example, the phrase *a fat dog bit John* can be represented by the propositions

(BITE, DOG, JOHN)

(QUALITY OF, DOG, FAT)

where BITE, DOG, JOHN, QUALITY OF, and FAT are concepts derived from the words in *a fat dog bit John*. In the first proposition, BITE is the predicate, while DOG and JOHN are arguments that stand in a particular relation to BITE. Specifically, DOG is the entity who BITES, and JOHN is the one who gets BITTEN. Similarly, in the second proposition, QUALITY OF is the predicate while DOG and FAT are arguments. Here, DOG is the entity that has the QUALITY and FAT is the QUALITY.

While propositional notation may appear cumbersome, the rationale behind it is simple: it is an attempt to represent the meaning of sentences in a formal and unequivocal way that allows research on sentence meaning to be done systematically. The meaning of sentences involves concepts that are distinct from the surface form of the text. For example, the phrase *John was*

bitten by a dog who was fat has a very different surface form than *a fat dog bit John* but approximately the same meaning. Within Kintsch and van Dijk's theory, to say that two sentences have the same meaning is to say that they have the same conceptual representation. Propositional notation is thus intended to be an approximation of this conceptual representation of sentences which is derived from, but is *not* identical with, their surface form.

As the propositions from each sentence are encoded, they are represented in memory as a coherent structure called a text base. The coherence Kintsch and van Dijk consider is the extent to which concepts (i.e., arguments) in a text are repeated. If two propositions have the same argument (like DOG in (BITE, DOG, JOHN) and (QUALITY OF, DOG, FAT)), the propositions are coherent because they contain the same argument. Thus, to say that a reader constructs a coherent text base is to say that the propositions extracted from the text are connected to each other in memory on the basis of argument repetition (see Bock and Brewer's chapter in this volume for other notions of coherence and its general importance to comprehension).

Kintsch and van Dijk's model of how a text base is constructed begins with the assumption that a reader is limited in how much can be kept in mind, or *short-term memory*, at any given time, and as a result a reader cannot comprehend a text all at once. Instead, the reader goes through the text in a series of processing *cycles*. For most texts, processing cycles correspond to comprehending a single sentence. Upon reading a sentence, the reader extracts the propositions that constitute the sentence's meaning, and connects the propositions of the sentence to each other on the basis of argument repetition. This processing of the propositions of a single sentence is what is known as a processing cycle. The propositions that are processed are said to be *in* the cycle.

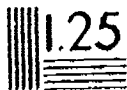
In order to illustrate cycling, the processing cycles for the propositions extracted from the first three sentences of the instructed-words story from the McKeown et al. (1983) study are presented in Figure 1. The cycles of the Kintsch and van Dijk model appear as the Normal model presented in left-most column. The propositions referred to in Figure 1 are presented in



28 25



22



proposition subordinate to the lead proposition (3) to be carried over into the next processing cycle. Thus, in cycle 2, proposition 2 remains the lead proposition, and the propositions extracted from the second sentence are connected to it. The leading edge strategy selects propositions 2 as the most superordinate and proposition 7 as the most recent proposition that is subordinate to the lead proposition to be carried over into cycle 3. In cycle 3, proposition 2 remains the lead proposition, and the propositions extracted from the third sentence are connected to it.

It sometimes happens that there is no argument overlap between the propositions carried over and newly encoded propositions. In this case, a proposition from an earlier cycle that does share an argument with one of the newly encoded propositions is retrieved from *long-term memory*. Retrieving propositions of a previous cycle from long-term memory is referred to as reinstating a proposition. Once a proposition is reinstated, a connection is made between the reinstated proposition and one of the newly encoded propositions. In summary, the text base of a text is constructed by organizing the propositions of each sentence into a coherence graph, and connecting these propositions to previously encoded propositions that are either carried over, or reinstated, from previous cycles.

On the basis of these processing assumptions, Kintsch and van Dijk's model predicts that the probability of a proposition being recalled increases as the number of cycles it is in increases. The specific relationship is mathematically described as $1 - (1 - p)^n$, where p equals the probability of a reader recalling a proposition that has appeared in a single cycle, and n equals the number of cycles in which the proposition has appeared. Using this mathematical expression for each proposition, the fit of the model can be tested against actual recall, by calculating an optimal value for p and computing the difference between the predicted and actual recall for all the propositions.

texts with unfamiliar words. Within the framework of Kintsch and van Dijk's theory, we will consider models of two ways that texts containing unfamiliar words may be comprehended. Both of these models share the same assumption about how propositions containing unfamiliar

words are encoded. According to Kintsch and van Dijk, comprehension involves connecting in memory newly encoded propositions (hereafter called new propositions) to previously encoded propositions (hereafter called old propositions). Encountering a proposition containing an unfamiliar word (hereafter called an unfamiliar proposition) could conceivably affect both its encoding and connection to new propositions. Since the Kintsch and van Dijk model describes comprehension after propositional encoding has taken place, it is not helpful in describing how the process of encoding itself might be affected. However, one can postulate that the product of the encoding process will be different for familiar and unfamiliar propositions. One possibility in this regard is that only a vague sense of the unfamiliar word's meaning will be encoded, resulting in an incomplete representation of the proposition. For example, if a reader is unfamiliar with the word *novice*, his or her representation of the proposition *the novice played the violin* may be akin to *someone played the violin*, rather than to *the beginner played the violin*.

Assuming that unfamiliar propositions are incompletely represented, there are several ways in which comprehension could be affected. One possibility is that a different text base might be constructed. A different text base would be constructed if cycling followed a *substitution* principle in which new propositions are connected only to familiar, and not to unfamiliar, propositions. The motivation for such a principle is that readers may attempt to process unfamiliar propositions as little as possible and as a result connect new information only to parts of the text that they have successfully understood (see Anderson and Freebody's (in press) *minimum effort principle*, for a similar notion).

Using Kintsch and van Dijk's model as a framework, the substitution principle can be represented by a Substitution model in which unfamiliar propositions are never carried over or reinstated into subsequent processing cycles, and familiar propositions are substituted in the place of unfamiliar propositions during cycling. The propositional structure created by this model differs from that created by the original Kintsch and van Dijk model. Unfamiliar propositions are connected to fewer propositions, and familiar propositions that otherwise would have been

connected to few propositions are connected to more propositions because they take the place of the unfamiliar propositions.

In order to illustrate the difference between the Substitution model and the original Kintsch and van Dijk model (hereafter called the Normal model), consider the processing cycles presented in Figure 1. Shown in this figure are the processing cycles of the Normal (column 1) and Substitution (column 2) models for the first three sentences of the instructed-words story used in the McKeown et al. study. In the first cycle of the Normal model, proposition 2, (ISA, SAM, NOVICE), is selected as the lead proposition. This proposition, along with proposition 3, (QUALIFY, 2, AT VIOLIN), is carried over into cycle 2. In the McKeown et al. study, *novice* was an unfamiliar word to the children in the control group. Therefore, in the Substitution model, even though proposition 2, (ISA, SAM, NOVICE), is in the lead position, it is not carried over, and instead, the familiar proposition 1, (EXIST, SAM), is carried over.

Shown at the bottom of Figure 1 is a summary of the number of cycles in which selected propositions appeared for the Normal and Substitution models. As can be seen from this summary, in the Substitution model, unfamiliar propositions enter fewer cycles, and the familiar propositions that substitute for the unfamiliar propositions enter more cycles than they do in the Normal model.

Another possibility of how unfamiliar words could affect comprehension is that the structure of the text base formed is not altered, but that the accessibility of unfamiliar propositions is reduced. This would occur if cycling followed a *suppression* principle in which new propositions are allowed to be connected to either familiar or unfamiliar propositions, but since unfamiliar propositions are assumed to be incompletely represented and therefore rather vague, the probability of their recall is suppressed. This is because it is only the incomplete representation of the unfamiliar proposition that is processed each time it is connected to a new proposition. The suppression principle assumes that repeated processing of an incomplete representation of the unfamiliar proposition will not enhance the probability of the unfamiliar

proposition appearing in recall as much as repeated processing of a complete representation, and in the extreme, recall of incomplete propositions may not be enhanced at all by repeated processing. To test the strong form of this approach, then, the suppression principle assumes that regardless of how many times the incomplete representation of an unfamiliar proposition is processed, the probability of the entire proposition being recalled will be the same as if it had been processed only once.

Using Kintsch and van Dijk's model as a framework, the suppression principle can be represented by a Suppression model in which unfamiliar propositions are allowed to be connected to new propositions during processing, but are given credit for appearing in only a single cycle. In order to illustrate the difference between the Suppression and Normal models, the processing cycles of the Suppression model for the first three sentences of the previously described instructed-words story are presented in column 3 of Figure 1. As is shown in this figure, the processing cycles of the Normal and Suppression models are virtually identical. The difference between the models lies in the fact that incomplete representations of unfamiliar propositions 2, 4, and 7 are processed in the Suppression model. As a result, the two models differ in the number of cycles for which these unfamiliar propositions receive credit, as shown at the bottom of the figure. In the Normal model, these propositions are given credit for all the cycles in which they appear. In contrast, in the Suppression model, these same propositions are given credit for appearing in only a single cycle and as a result are predicted to be recalled less well.

The difference between the Suppression and Substitution models is also shown in Figure 1. As shown at the bottom of Figure 1, the Substitution and Suppression models both give credit to unfamiliar propositions for being in only a single cycle. The two models differ in that familiar propositions 1, 3, and 6 appear in more cycles in the Substitution model than they do in the Suppression model. In the Substitution model, unfamiliar propositions appear in only a single cycle because they are not used in processing subsequent to encoding, and selected familiar propositions appear in additional cycles because they are substituted in place of the unfamiliar

propositions during carryovers. In the Suppression model, unfamiliar propositions likewise appear in only a single cycle, because an incomplete representation which does not benefit the unfamiliar proposition's recall, is carried over during cycling. However, because the incomplete representations are used, there is no substitution, and consequently, the processing of familiar propositions is not affected.

texts with recently taught words. Kintsch and van Dijk's model can also be used as a framework for describing comprehension of texts containing words recently taught. We will consider here two ways in which such texts may be comprehended.

When a reader encounters a proposition containing a word about which the reader has received recent prior instruction (hereafter called an instructed proposition), there are several possibilities about how processing may be affected. One possibility is that processing will follow a normal processing principle in which instructed propositions are treated in the same way as are familiar propositions. Such a principle is assumed by most studies examining the effects of vocabulary instruction on comprehension (e.g., Beck et al., 1982; Kameenui et al., 1982; McKeown et al., 1983).

Using Kintsch and van Dijk's model as a framework, the normal processing principle can be represented by a Normal model in which instructed and familiar propositions are processed in the same way. Such a model is the original version of Kintsch and van Dijk's (1978) model of comprehension, which makes no distinction between instructed and uninstructed propositions.

A second possibility of how instructed words may affect processing is that processing will follow a remind principle in which encountering an instructed word will remind the reader that that word was the object of prior instruction. A number of researchers have demonstrated that remembering the context in which a fact was learned can enhance one's memory of the learned fact (e.g., Jacoby & Craik, 1979; Jacoby & Dallas, 1981; McKoon & Ratcliff, 1979). Within this same vein, the remind principle assumes that recognizing a word as an instructed word involves making a connection between the instructed proposition and the reader's representation of his or

her instructional experience. Thus, instructed propositions are initially processed twice: once during encoding, and a second time as a connection is made between the instructed proposition and the reader's representation of the prior instructional experience.

Using Kintsch and van Dijk's model as a framework, the remind principle can be represented by a Remind model in which instructed propositions appear in their initial cycle twice. One of these is due to the proposition being encoded into the coherence graph. The other is due to a connection that is made between the proposition and the reader's representation of the prior instructional experience. This remind connection does not become part of, or affect in any way, the coherence graph of the text. Instead, it is processing during which the reader recalls from memory the learning context of the word. Even though being reminded of a word's learning context is not part of the comprehension process per se, it is additional processing and as a result might be expected to augment the probability of recall much in the same way as a text reinstatement would. Thus, the Remind model expands the Kintsch and van Dijk model to account for not only the effects of text processing, but also the effects of connections that are made to episodic memory in parallel to text processing.

In order to illustrate the difference between the Remind and Normal models, the processing cycles of the Remind model for the first three sentences of the previously described instructed-words story are presented in column 4 of Figure 1. As shown in this figure, the coherence graphs constructed for the cycles of the Normal and the Remind models are identical. The two models differ in that in cycle 1 of the Remind model, instructed proposition 2, (ISA, SAM, NOVICE), appears twice, once due to its initial encoding into the coherence graph, and once due to its connection to the reader's representation of the prior instructional experience. In cycle 1 of the Normal model, instructed proposition 2 appears only once due to being encoded into the coherence graph. Similarly, in cycle 2 of the Remind model, instructed proposition 4, (GIVE, VIRTUOSO, CONCERTS), and instructed proposition 7, (QUALITY OF, SAM, AMBITIOUS), appear twice. The cycles for selected instructed and familiar propositions shown at the bottom of

Figure 1 show that the difference in the number of cycles credited for the propositions of the Normal and Remind models is that the instructed propositions receive credit for being in an additional cycle in the Remind model as compared to the credited cycles of the Normal model.

results. In order to assess how well the different models predicted recall, a computer program was used that calculated a value for p that minimized the difference between the predicted and actual recall (Chandler, 1965). The fit of each model then was obtained by first summing the differences between the predicted and actual recall scores for each proposition and standardizing this sum by taking the square root of its square divided by the number of propositions. This fit score is referred to as the root-mean-square-deviation (RMSD) score. Since RMSD scores reflect deviation from prediction, the *lower* the score, the *better* the fit. The RMSD scores for the four models are presented in Figure 2.

Insert Figure 2 about here

As shown in Figure 2, the Suppression model best predicted the pattern of recall of the three story situations in which the children were unfamiliar with some of the words. Moreover, as shown in Figure 3, the Suppression model best predicted the recall of these stories by both skilled and less-skilled readers. These results suggest that when both skilled and less-skilled readers encounter unfamiliar propositions, they attempt to treat them in the same way as they do familiar propositions rather than skipping over them. Thus, the adverse effects of unfamiliar propositions on text recall are likely due only to making some parts of the text base inaccessible as suggested by the Suppression model, rather than to the construction of a different text base as suggested by the Substitution model.

Insert Figure 3 about here

As is also shown in Figure 2, the Remind model best predicted the pattern of recall of the story situation in which the children had been taught some of the words. Moreover, as shown in Figure 3, the Remind model best predicted the recall of these stories by both skilled and less-skilled readers. These results suggest that when both skilled and less-skilled readers encountered propositions containing instructed words, the fact that the target words had been the object of previous classroom instruction was called to mind, which entailed additional processing that enhanced recall.

implications

implications for theories of processing. The fact that the Suppression model, rather than the Substitution model, best predicts the pattern of recall of texts containing unfamiliar words suggests that readers may attempt to minimize the effect that difficulties encountered during the processing of words has on comprehension. Specifically, even though propositions containing unfamiliar words are likely to be incompletely represented, these results suggest that readers attempt to construct the same text base as would be constructed if the unfamiliar propositions were familiar. Thus during cycling, they do not seem to substitute familiar propositions for unfamiliar ones. Rather, they attempt to use the incomplete representations they construct for unfamiliar propositions in subsequent processing. The fact that the Suppression model better predicted the pattern of recall than did the Normal model indicates that although the structure of the text base formed was unaltered, the unfamiliar propositions in the text base were relatively inaccessible during recall.

The fact that the Remind model best predicts the pattern of recall of the story containing instructed words suggests that children can be familiar with words in differing ways which can have various effects on comprehension. The instruction children receive may make the context in which the word was learned so salient that it is called to mind whenever an instructed word is encountered. The result is that words learned through the instruction may be treated differently than common words the children already knew. For example, a child may be very familiar with

the word *bird*, but have no recollection of when or where the word was learned. Such a word would not disrupt processing, but neither would it initiate additional processing. In contrast, if a child has recently been taught *miser* through direct instruction, encountering the word in a text could initiate the additional processing of calling to mind the learning context, which in turn would increase its probability of being recalled. It is quite possible that the salience of the learning context may diminish in a few weeks, especially as subsequent instruction on new words is encountered. However, without additional data, one can only speculate about how permanent or transitory this biasing effect of the learning context is.

The fact that the same pattern of results was obtained for skilled and less-skilled readers suggests that skilled and less-skilled readers adopt similar strategies in adapting to variation in word knowledge. There are many well documented differences between skilled and less-skilled readers. Less skilled readers read words more slowly (Perfetti & Hogaboam, 1975), with less semantic activation (Jackson & McClelland, 1979), and less verbatim memory, (Goldman, Hogaboam, Bell, & Perfetti, 1980), than do skilled readers. However, they do not seem to differ in the nature of their susceptibility to the effects of variation in word knowledge. Evidently knowing words is a skill that is critical to comprehension regardless of one's reading ability. (Of course, less-skilled readers may be more affected by this susceptibility in absolute terms than are skilled readers simply because less-skilled readers know fewer words on the average than skilled readers (Davis, 1944, 1968; Singer, 1965; Thurstone, 1946)).

In summary, lexical knowledge has a wide range of effects on comprehension. Lexical items that are unfamiliar give rise to the construction of incomplete propositions that apparently do not result in the construction of a different text base by disrupting cycling, but do make recall of the propositions containing the unfamiliar words difficult. Lexical items that have recently been learned through direct instruction likewise do not appear to affect the maintenance of coherence during processing, but do give rise to parallel processing in which the learning context of the instructed words is called to mind, which in turn improves the recall of propositions containing

the instructed words.

implications for education. There are implications of this work for both contextual and direct instructional techniques of teaching vocabulary. Many commercial basal reading programs assume that the meanings of new words can be derived from context when they are encountered in a story (Beck, McKeown, McCaslin & Burkes, 1979). The fact that the Suppression rather than the Substitution model best predicted recall is consistent with the assumption of contextual methods that readers attempt to use what meaning they glean from context during processing rather than simply ignoring the unfamiliar word. However, the fact that recall was impaired by the presence of unfamiliar words serves as a cautionary note to an overreliance on contextual methods. Since the presence of unfamiliar words can impair comprehension, it seems wise to make sure that the contexts in which new words appear are "pedagogical" in the sense that they be created with the intent to lead the reader to a specific, correct meaning of the word. At present, story contexts in basal readers are not so constructed (Beck, McKeown, & McCaslin, 1983).

The fact that the Remind model best predicted the recall of the instructed-words story by the experimental children reflects the fact that children were more likely to remember the parts of the story containing instructed words than other parts. This suggests that learning words through well designed, direct instruction may make the instructed words particularly salient and result in a greater ability (or at least a greater proclivity) to use the instructed words during comprehension than many familiar words. Such a possibility underscores the fact that the tendency to use a word during comprehension is one of the many dimensions of "knowing" a word that exists (cf. Beck et al., 1982; Curtis & Glaser, 1983). If we are to succeed in understanding and controlling the difficulty children have with reading, it will be important to better understand not only how comprehension is affected by word unfamiliarity, but also how it is affected by different types of vocabulary instruction through which new words are learned.

implications for modeling. The focus of the work reported in this chapter has been on the

construction and evaluation of models of text comprehension. Given the time and effort involved in modeling, it is appropriate to compare its strengths and limitations.

The strength of modeling is that it expresses global principles about processing in an explicit form that can be evaluated. For example, a general notion like *unfamiliar propositions are avoided during processing* can be translated into a specific model (in this case the Substitution model) from which recall predictions for each proposition can be derived. The explicitness that comes from translating a general principle into a particular model has a number of benefits. First, it enables alternate principles to be tested. As long as principles are only globally described, it is difficult to assess the relative merit of competing principles. Committing the principles to explicit models provides the type of detail needed to evaluate which principle best accounts for observed performance.

A second benefit of committing general principles to explicit models is that often unexpected effects of the principles are discovered. For example, when the Substitution model was first constructed from the substitution principle, it was an unexpected result that the text base would be disrupted to the extent it was. This realization suggested that an alternative principle would be one that did not alter the text base.

The limitation of modeling is that the only theoretical explanations that can be considered are those that are employed by the particular model chosen. In the present study, basing the models on Kintsch and van Dijk's (1978) theory of text comprehension limits the nature accounts given in several ways. First, it necessitates the use of propositions rather than words as the unit of analysis, which in turn may result in over-estimating the effect of unfamiliar and instructed words. It may be that encountering an unfamiliar words impairs the recall of only that word rather than of the entire proposition in which the word is contained. However, within Kintsch and van Dijk's theory, the recall of individual words cannot be represented and consequently, this alternative cannot be tested.

A somewhat more serious limitation involves alternative explanations for the observed bias

in recalling instructed propositions. The explanation used in the Remind model is that instructed propositions receive additional processing as a result of their being connected to representations of the instructional contexts. There are, however, a number of alternative candidates that do not fit easily into the Kintsch and van Dijk framework. One alternative account is that during encoding the fact that the instructed propositions contains recently learned words may make them distinctive from the rest of the propositions and as a result they would be more salient in the representation of the text and consequently are more likely to be recalled (e.g., Bransford, Franks, Morris, & Stein, 1979; Nelson, 1979; Eysenck, 1979). This account could not be represented within the Kintsch and van Dijk framework without changing the basic assumption that the probability of a proposition being recalled is due to the number of times it enters a processing cycle rather than to how it is encoded. (It turns out, however, that the bias to recall instructed propositions is probably not due to their distinctiveness in comparison to uninstructed propositions. This is because a simple notion of distinctiveness is unable to account for why unfamiliar propositions, which are also distinctive in comparison to familiar propositions, are *less* likely to be recalled.)

A another alternative account of the bias to recall instructed propositions that similarly involves the *manner* in which instructional propositions are encoded is that instructed propositions may be encoded more completely than other propositions and as a result may be more accessible during recall. As with the distinctiveness of encoding alternative, this completeness of encoding alternative could not be represented with the Kintsch and van Dijk framework without changing basic assumptions. (Again, it turns out that the bias to recall instructed propositions is probably not due to their being more accessible during recall. McKeown et al. (1983) report that children were no faster in categorizing instructed words than they were common words which suggests that the instructed words were not more accessible than common uninstructed words.)

A third alternative account of the bias to recall instructed propositions that does not easily

fit into Kintsch and van Dijk's framework is that readers may generate more elaborations for instructed than for uninstructed propositions. For example, upon reading *Sam was a novice at playing the violin*, readers probably generate a number of specific details that are consistent with, but not necessarily implied by, the text. For instance, it might be inferred that Sam was just beginning to take violin lessons, that the violin squeaked when he played it, and that the violin he owned was new and shiny. Since Kintsch and van Dijk's model deals only with that part of a text representation that is directly derived from the text, there is no simple way in which reader-generated elaborations can be predicted or even represented. (The bias to recall instructed propositions is also unlikely to be due to elaborations. A critical assumption to elaboration theories such as that of Anderson and Reder (1979) is that it is easier to generate elaborations about material that is familiar than about material that is unfamiliar. There were many familiar words in the uninstructed propositions about which the children knew a great deal more than the recently instructed words. Therefore, according to most elaboration theories, propositions containing only familiar propositions should have been better recalled than those containing recently instructed propositions.)

A final limitation of modeling is that often there is in principle an unlimited number of models that could be considered. For example, in the present study, a model of the comprehension of texts with instructed propositions could have been constructed from a biasing principle in which instructed propositions are always carried over into the next cycle regardless of whether they are part of the leading edge. A model based upon this principle would have been interesting to consider because it postulates that the improvement in recall for instructed propositions actually alters the text base that is constructed. There are doubtless other models that would have been equally interesting to consider. The fact that there are an indefinite number of models that can be constructed for a given phenomenon underscores the fact that modeling is only as useful as are the questions for which the models are designed to answer. Modeling is a powerful tool to decide between two competing accounts of processing. It does not

guarantee, however, that the accounts being evaluated are of theoretical or practical interest. There is a saying in computer science that goes like this: "Garbage in -- garbage out." The point of the saying is that there is nothing *magic* about using a computer as a tool to solve a problem. The solution you get is only as good as the way you formulate the problem and the data you give the computer to process. This same thing is true of modeling as a tool. It is only as good as the theories being modeled.

references

- Anderson, J. R., & Reder, L. M. "An Elaborative Processing Explanation of Depth of Processing." In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of Processing and Human Memory*. Hillsdale, NJ: Lawrence Erlbaum, 1979.
- Anderson, R. C., & Biddle, W. B. "On Asking People Questions About What They Are Reading." In G. H. Bower (Ed.), *Psychology of Learning and Motivation*. New York: Academic Press, 1975.
- Anderson, R. C., & Freebody, P. "Reading Comprehension and the Assessment and Acquisition of Word Knowledge." In B. Hutson (Ed.), *Advances in Reading/Language Research: A Research Annual*. Greenwich, CT: JAI Press, in press.
- Beck, I. L., McKeown, M. G., & McCaslin, E. S. "Vocabulary Development: All Contexts Are Not Created Equal." *Elementary School Journal*, 1983, 83 (3), 177-181.
- Beck, I. L., McKeown, M. G., McCaslin, E. S., & Burkes, A. M. *Instructional Dimensions that May Affect Reading Comprehension: Examples from Two Commercial Reading Programs*. Technical Report No. 1979/20. Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center, 1979.
- Beck, I. L., Perfetti, C. A., & McKeown, M. G. "The Effects of Long-Term Vocabulary Instruction on Lexical Access and Reading Comprehension." *Journal of Educational Psychology*, 1982, 74 (4), 506-521.
- Bower, G. H. "Experiments on Story Comprehension and Recall." *Discourse Processes*, 1978, 1 (3), 211-231.
- Bransford, J. D., Franks, J. J., Morris, C. D., Stein, B. S. "Some General Constraints on Learning and Memory Research." In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of Processing and Human Memory*. Hillsdale, NJ: Lawrence Erlbaum, 1979.
- Eysenck, M. W. "Depth, Elaboration, and Distinctiveness." In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of Processing and Human Memory*. Hillsdale, NJ: Lawrence Erlbaum, 1979.

- Chandler, P. J. *Subroutine STEPTT: An Algorithm that Finds the Values of the Parameters Which Minimize a Given Continuous Function*. Bloomington, IN: Indiana University.
- Curtis M. E., & Glaser, R. "Reading theory and the assessment of reading achievement." *Journal of Educational Measurement*, 1983, 20 (2), 133-147.
- Davis, F. B. "Fundamental Factors of Comprehension in Reading." *Psychometrika*, 1944, 9(2), 185-197.
- Davis, F. B. "Research in Comprehension in Reading." *Reading Research Quarterly*, 1968, 3(4), 499-545.
- Draper, A. G., & Moeller, G. H. "We Think with Words (Therefore, to Improve Thinking, Teach Vocabulary)." *Phi Delta Kappan*, 1971, 52 (4), 482-484.
- Goldman, S. R., Hogaboam, T. W., Bell, L. C., & Perfetti, C. A. "Short-Term Retention of Discourse During Reading." *Journal of Educational Psychology*, 1980, 72 (5), 647-655.
- Jackson, M. D., & McClelland, J. J. "Processing Determinants of Reading Speed." *Journal of Experimental Psychology: General*, 1979, 108 (2), 151-181.
- Jacoby, L. L., & Craik, F. I. M. "Effects of Elaboration of Processing at Encoding and Retrieval: Trace Distinctiveness and Recovery of Initial Context." In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of Processing and Human Memory*. Hillsdale, NJ: Lawrence Erlbaum, 1979.
- Jacoby, L. L., & Dallas, M. "On the Relationship Between Autobiographical Memory and Perceptual Learning." *Journal of Experimental Psychology: General*, 1981, 110(3), 306-340.
- Jenkins, J. R., Pany, D., & Schreck, J. *Vocabulary and Reading Comprehension: Instructional Effects*. Technical Report No. 100. Champaign, IL: University of Illinois, Center for the Study of Reading, 1978.
- Just, M. A., & Carpenter, P. A. "A Theory of Reading: From Eye Fixations to Comprehension." *Psychological Review*, 1980, 87 (4), 329-354.
- Kameenui, E. J., Carnine, D. W., & Freschi, R. "Effects of Text Construction and Instructional

- Procedures for Teaching Word Meanings on Comprehension and Recall." *Reading Research Quarterly*, 1982, 17(3), 367-388.
- Kintsch, W. *The Representation of Meaning in Memory*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1974.
- Kintsch, W., & van Dijk, T. A. "Toward a Model of Text Comprehension and Production." *Psychological Review*, 1978, 85(5), 363-394.
- Lehnert, W. G. *The Process of Question Answering*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1978.
- Lesgold, A. M., & Perfetti, C. A. (Eds.). *Interactive Processes in Reading*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1981.
- McKeown, M. G., Beck, I. L., Omanson, R. C., & Perfetti, C. A. "The Effects of Long-Term Vocabulary Instruction on Reading Comprehension: A Replication." *Journal of Reading Behavior*, 1983, 15(1), 3-18.
- McKoon, G., & Ratcliff, R. "Priming in Episodic and Semantic Memory." *Journal of Verbal Learning and Verbal Behavior*, 1979, 18(4), 463-481.
- Nelson, D. L. "Remembering Pictures and Words: Appearance, Significance, and Name." In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of Processing and Human Memory*. Hillsdale, NJ: Lawrence Erlbaum, 1979.
- Omanson, R. C., Beck, I. L., McKeown, M. G., & Perfetti, C. A. *Comprehension of Texts with Unfamiliar Versus Recently Taught Words: An Assessment of Alternative Models*. Unpublished manuscript. Pittsburgh, PA: Learning Research and Development Center, University of Pittsburgh, 1984.
- Perfetti, C. A., & Hogaboam, T. W. "The Relationship Between Single Word Decoding and Reading Comprehension Skill." *Journal of Educational Psychology*, 1975, 67(4), 461-469.
- Rumelhart, D. "Toward an Interactive Model of Reading." In S. Dornic & P. M. A. Rabbitt (Eds.), *Attention and Performance VI*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1977.

- Schank, R. C. "The Structure of Episodes in Memory." In D. G. Bobrow & A. Collins (Eds.), *Representation and Understanding: Studies in Cognitive Science*. New York: Academic Press, 1975.
- Singer, H. "A Developmental Model of Speed of Reading in Grades 3 Through 6." *Reading Research Quarterly*, 1985, 1(1), 29-49.
- Stein, N. L., & Glenn, C. G. "An Analysis of Story Comprehension in Elementary School Children." In R. Freedle (Ed.), *Advances in Discourse Processes*. Volume 2. Hillsdale, NJ: Lawrence Erlbaum Associates, 1979.
- Thurstone, L. L. "A Note on a Reanalysis of Davis' Reading Tests." *Psychometrika*, 1946, 11(2), 185-188.
- Tuinman, J. J., & Brady, M. E. "How Does Vocabulary Account for Variance on Reading Comprehension Tests? A Preliminary Instructional Analysis." In P. Nacke (Ed.), *Twenty-Third National Reading Conference Yearbook*. Clemson: The National Reading Conference, 1974.

Richard C. Omanson is a postdoctoral fellow at the Learning Research and Development Center at the University of Pittsburgh. He received his Ph.D. in child psychology from the University of Minnesota.

Table 1

Prototypical Recalls of the Instructed-Words
 Story for the Experimental and Control Groups

Experimental	Control
<p>There was a man named Sam. Sam was a NOVICE at playing the violin. Sam gave a concert. There was a woman. There was a woman in the audience. The OBESE woman said something to an ACQUAINTANCE. There was something EDIBLE. The woman began to DEVOUR everything. Sam asked if the woman couldn't FAST. Sam THRUST his violin on the stage and began to TRUDGE away. There was an ALLY. He SEIZED Sam and suggested that they all play music. Some people began to play instruments. Everyone felt PLACID.</p>	<p>There was a man named Sam. Sam played the violin. He rented a hall, invited some people, and gave a concert. He began to play some notes. There was a woman in the audience. Some people began to play instruments. Everyone went home feeling PLACID.</p>

Note. Instructed words are capitalized.

Table 2

- Propositions Contained in the First Three
Cycles of the Instructed-Words Story

Propositions	Text
Cycle 1	
1. (EXIST, SAM)	Sam was a NOVICE at playing the violin.
2. (ISA, SAM, NOVICE)	
3. (QUALIFY, 2, AT VIOLIN)	
Cycle 2	
4. (GIVE, VIRTUOSO, CONCERTS)	He knew that usually a VIRTUOSO gave
5. (QUALIFY, 4, USUALLY)	concerts, but he was AMBITIOUS.
6. (KNOW, SAM, 4)	
7. (QUALITY OF, SAM AMBITIOUS)	
8. (CONTRAST, 4, 7)	
Cycle 3	
9. (RENT, SAM, HALL)	So one day Sam rented a music hall and
10. (QUALIFY, HALL, MUSIC)	invited some people to hear him play.
11. (TIME, 9, ONE DAY)	
12. (CAUSE, 9, 7)	
13. (INVITE, SAM, PEOPLE)	
14. (PLAY, SAM)	
15. (HEAR, PEOPLE, 14)	
16. (PURPOSE, 13, 14, 15)	

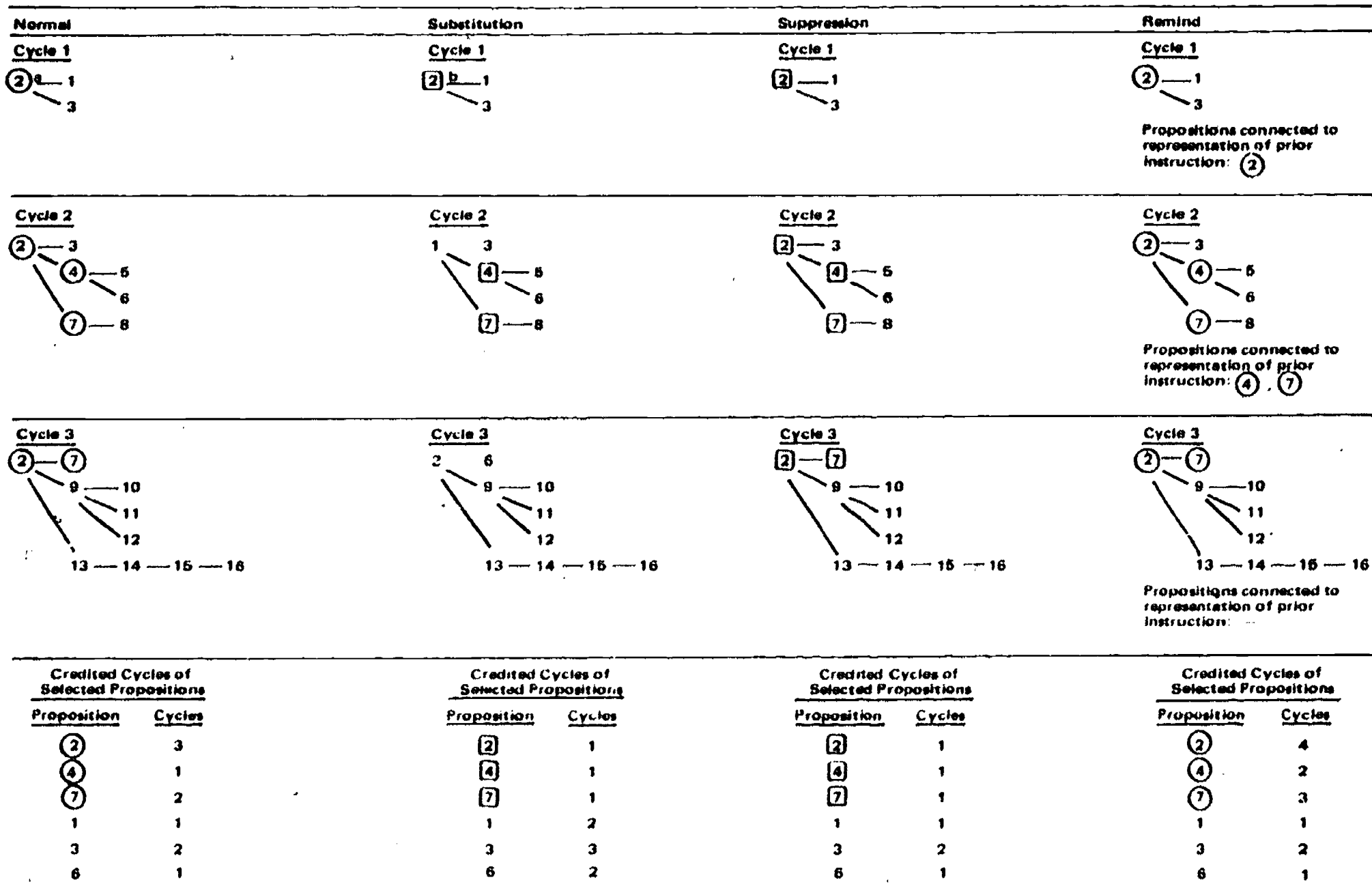
Note. Words contained in the text that were either unfamiliar or instructed are capitalized.

figure captions

Figure 1. First three cycles of the Normal, Substitution, Suppression, and Remind models for the instructed-words story.

Figure 2. RMSD scores for models of the instructed- and uninstructed-words stories.

Figure 3. RMSD scores for models of the instructed- and uninstructed-words stories for skilled and less-skilled readers.



^a Propositions assumed by model to be instructed are circled.

^b Propositions assumed by model to be unfamiliar and incompletely represented are boxed.

Figure 1

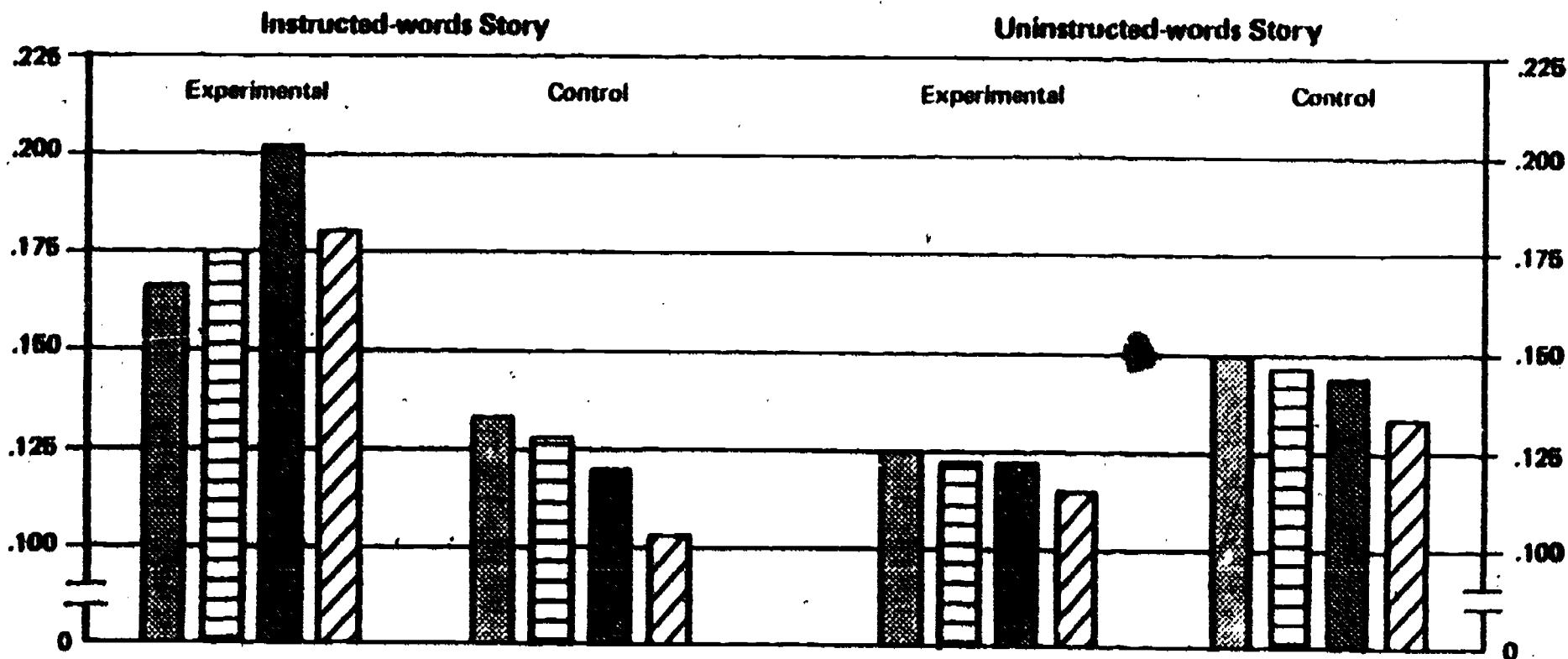
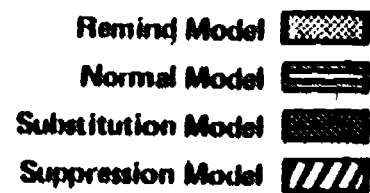


Figure 2

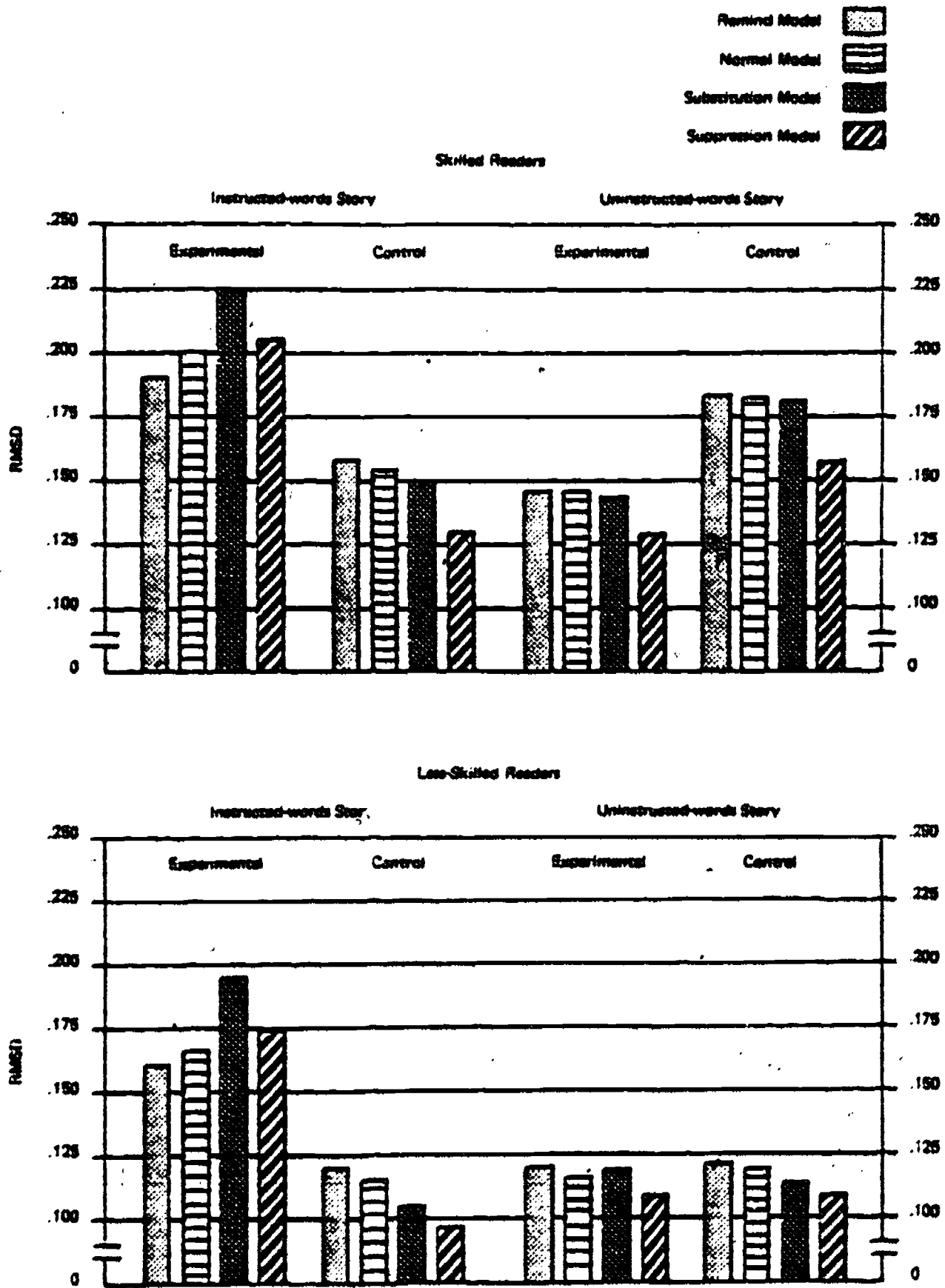


Figure 3