DOCUMENT RESUME

ED 094 317		CS 001 182
AUTHOR	Lesgold, Alan M.	
TITLE	The Role of Mental Imagery in Text Comprehensicn: Preliminary Studies.	
INSTITUTION	Pittsburgh Univ., Pa. Learning Research and Development Center.	
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.	
PUB DATE	74	
NOTE	71p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, April 15-19, 1974)	
EDRS PRICE	MF-$0.75 HC-$3.15 PLUS POSTAGE	
DESCRIPTORS	Comprehension; *Imagery; *Learning; Learning Processes; *Memory; *Reading Comprehension; Symbolic Learning	

ABSTRACT

Studies of the effects of imagery instructions and other imagery induction on prose learning are included in this progress report. Five experiments were conducted to establish a developmental trend for ability to use imagery-related mediation in whole passage learning. Observations indicate that six year olds are inhibited in paraphrase recall of heard prose when they are asked to illustrate the passages with cut-out figures, but that nine year olds show no imagery effect. After an extensive training session, nine year olds show a positive effect. Adults show a spontaneous positive effect. The results do not appear to require imaginal-visual-spatial explanations. Information gathered in observations suggests the hypothesis that differential attention to and structuring of working-memory content imagery instructions facilitate prose learning. Further research involves various experiments designed to test this hypothesis. (Author/JM)

THE ROLE OF MENTAL IMAGERY IN TEXT COMPREHENSION: PRELIMINARY STUDIES

Alan M. Lesgold

Mary E. Curtis

Hildrene De Good

Roberta M. Golinkoff

Claire McCormick

and

Joseph Shimron

Learning Research and Development Center

University of Pittsburgh

1974

This paper reports research conducted with the support of the Learning Research and Development Center, University of Pittsburgh, through funds provided by the National Institute of Education. The opinions expressed are solely those of the authors and should not be taken as indicating the policies of either organization. The students and staff of Madison, Weil, and Frick Schools of the Pittsburgh Board of Education, the Falk School, and St. Agnes School of the Diocese of Pittsburgh provided assistance upon which this work depended. Charles Perfetti, James Voss, and Margaret Wang provided helpful comments on an earlier draft.

Abstract

This paper is a progress report on studies of the effects of imagery instructions and other imagery inductions on prose learning. Initially, a number of theoretical issues are discussed followed by the presentation of a series of five experiments in which a developmental trend for ability to use imagery-related mediation in whole passage learning is established. Observations indicate that six-year-olds are inhibited in paraphrase recall of heard prose when they are asked to illustrate the passages with cut-out figures, but that nine-year-olds show no imagery effect. After an extensive training session, nine-year-olds show a positive effect. Adults show a spontaneous positive effect. The results do not appear to require imaginal-visual-spatial explanations. From information gathered in observations, a hypothesis is proposed: Differential attention to and structuring of working-memory content under imagery instructions facilitate prose learning. Various experiments designed to test this hypothesis, which are now in progress or under development, are described.

Table of Contents

Table of Contents (Continued)

THE ROLE OF MENTAL IMAGERY IN TEXT COMPREHENSION: PRELIMINARY STUDIES

Alan M. Lesgold,[1] Mary E. Curtis, Hildrene De Good, Roberta M. Golinkoff, Claire McCormick, and Joseph Shimron

Learning Research and Development Center
University of Pittsburgh

I. Theoretical Rationale

A. Definitions and Purpose

This is a progress report on the effect of mental imagery instructions and related tasks on prose learning. It deals with two basic questions: (a) does "mental imagery" effect prose learning; and (b) is there a developmental trend in those effects? First, various hypotheses that predict imagery task/instruction effects are presented followed by a number of experiments that test the existence of such effects and show the developmental trend. Finally, there is a brief description of plans for future work.

In all of the studies reported here, prose comprehension is measured by the amount of content recalled from passages that have been read. This paraphrase recall measure is described in detail in Section II. A. 1. It is often argued that memory measures are inappropriate for measuring comprehension. We choose not to follow this line of argument. In this report we state the measures that we have used;

---

[1]Also in Department of Psychology, University of Pittsburgh.

it is left to others to determine whether or not our work is concerned with comprehension.

Similarly, we have not taken a particularly constrained view of what we mean by mental imagery. All of our imagery manipulations have been instructional manipulations. We tell subjects to have a picture in their minds of what is going on in the passages they read, or to read the passage to extract the information necessary to either draw a cartoon or construct a picture from cut-outs. Section III will deal with possible changes in the kinds of imagery manipulations we use.

B. Why Imagery Mediation of Prose Learning is Expected

1. Extending Paired Associate Results

Recent researchers have demonstrated that associative learning is greatly facilitated by instructions to form mental images in which the referents of word pairs to be learned are seen in some kind of vivid interaction (see Bower, 1972; Paivio, 1971). The effects are massive--it is not uncommon that the amount one recalls doubles in such experiments (Bower, 1972, p. 67). Research in paired associate imagery has continued in three directions. First, the extent to which imagery processing involves cognitive operations that are distinct from verbal processing is being explored. Second, research on the possibility of extending the "magic" of imagery to such real world tasks as reading has burgeoned. Third, researchers are attempting to define the specific cognitive processes that underlie the imagery effect; they are studying the development of these processes. In this section we will examine the empirical evidence that suggests utility for imagery instructions in prose learning. Then we will consider the most elementary theoretical explanation for imagery effects--meaningfulness--and show it to be inadequate. Subsequent sections will examine more tenable explanations.

There is a body of experimental evidence that suggests that imagery instruction effects can be extended from paired associate learning to prose passage learning. Richard Anderson (1971) has shown a facilitative effect for imagery instructions compared to rote repetition instructions in intentional learning of unrelated sentences. In another experiment Anderson (R. Anderson & Hidde, 1971) showed that incidental learning by subjects who rated sentence imageability was three times as high as that of subjects who rated sentence pronounceability. This latter experiment is open to the argument that pronounceability rating may be more distracting than simple learning. The former experiment compares imagery mediation to a more suitable control. In practical terms, it would be nice to see a comparison of imagery to self-constructed strategies. This would show whether there is any practical import to imagery instruction at the sentence level. Further, the measures of recall might be broadened beyond the word-by-word scoring of Anderson.

At the level of prose learning, the data become murky, but three studies show evidence of an imagery facilitation effect. Anderson and Kulhavy (1972) gave high school age subjects instructions to form mental images as they read prose. These subjects' performance was compared to that of subjects told only to "read carefully." No differences were found between groups. However, within both groups, subjects who reported using mental imagery performed better. The problem, though, is that such post hoc reports may be biased by the relative feelings of success or failure that the subject has in his criterion performance.

A second study (Matz & Rohwer, 1971) found that pictorial presentation of information along with sentences expressing that information enabled low-SES, black fourth graders to perform at the level of middle-SES, white fourth graders; in the absence of the pictures, the low-SES blacks do worse than the middle-SES whites. Levin (1973) has shown that this pictorial presentation is particularly helpful for poor readers who have a vocabulary or decoding deficit. Poor comprehenders with adequate decoding skills and good comprehenders benefit instead from imagery instructions. Both the Matz and Rohwer and the Levin studies only approximate real reading situations. This is because the sentences are presented for study one at a time.

Levin's study does, however, demonstrate a difference between generating an image and being provided with a picture. Good comprehenders benefit primarily from imagery that they generate themselves, just as adults benefit most from self-generated (versus experimenter-provided) verbal mediation (Bobrow & Bower, 1969). Young children (5-11 years), on the other hand, benefit mainly from experimenter-provided verbal mediators (see Rohwer, in press). Also, poor readers with poor decoding skills benefit from pictures in Levin's study but not from imagery instructions. We do not think that these results run counter to the basic principle that self-generated mediation is better than experimenter-provided mediation. When we consider developmental trends in Section I. C., we will consider the possibility that experimenter mediation is successful in exactly those cases when the subject does not have adequate ability to generate his own mediator.

To summarize, several experiments suggest a facilitative effect of imagery on prose learning. However, all of these demonstrations have involved sentence-by-sentence presentation. What remains unanswered is whether imagery instructions will facilitate performance in uninterrupted reading of a whole passage. The present experiments are all conducted at the whole passage level. The importance of this distinction between sentence-by-sentence and whole-passage processing can be better understood by asking again why we expect imagery to facilitate prose retention.

Expectations for imagery facilitation of text rest mainly on naive-empirical grounds. It worked for paired associates and for sentences, so why not shoot upward one more level. In fact, that upward step has been minimal. Prose retention in the R. Anderson (1971; Anderson & Hidde, 1971; Anderson & Kulhavy, 1972) and in the Levin (1973) studies was measured in a very associative fashion. Either recall of the predicate of a sentence was requested, given the subject as a probe, or questions were asked that usually amounted to probes for an association to some fragment of the text. The presentation unit has remained the sentence, though it may be preceded or followed by other related sentences that are displayed separately.

Thus, the demonstrations of imaginal facilitation of prose learning are weak. The most common theoretical explanation is even weaker. Either the laboratory paired associate task is equated with comprehension and learning in a natural setting, or the simple association is presumed to be the atomic component of all learning. The former position is an act of faith, while the latter is in conflict with much of our knowledge of language comprehension and production (see the chapter on associationism in J. Anderson & Bower, 1973). Even if we accept the thesis that knowledge is a bundle of associations, the value of elaborative mediators for prose is unclear. Elaborative mediation is assumed to facilitate prose learning because it enables the subject to find some basic meaningful property or event common to the paired words' referents (Rohwer, in press) and to derive from that central relationship the simplest and strongest primitive encoding for the pair (Bower, 1972).

However, sentences are <u>already meaningful</u>. A relationship between subject and object is given by the sentence. Perhaps random sentences, that have only generic meaning, require further elaboration to add some common meaning that holds them together. However, in the limiting case of conventional communication in a natural setting, even the intersentence association or meaningfulness should already be there. If imagery mediation works by rendering the meaningless meaningful, then it should have less benefit as the materials become more like natural communications.

The above arguments are not meant to imply that there is not a common explanation for imagery effects in paired associate learning and prose learning. Rather, we wish to point out the weaknesses of current explanations of paired associate mediation. Perhaps a critical examination of imagery effect in prose learning will provide better theory for understanding imagery effects in simple laboratory tasks. Before doing that, we briefly consider several other possible explanations.

2. <u>Dual-Process Models</u>

Another standard explanation for imagery mediation effects argues that verbal and imaginal processes differ in many ways, and that the ability to use both processes is useful in learning verbal material. Imagery may be effective as a mediation device because it is a way of representing certain information in text more simply. One kind of representation ability that imagery may add is a calculus of spatial relationships that overcomes the strongly sequential pattern of verbal representations (Paivio, 1971). For example, it is simpler to have a spatial image of the locations of the pieces on a chessboard than to have a written description.

This point can be recapitulated in terms of the student's mental activities. Perhaps the student who comprehends text effectively is actively generating alternative verbal and nonverbal representations of the to-be-learned information structure. He may then focus his attention on the representations that are simplest, checking them against the text to see that they are complete.

The extent to which imaginal thought differs from verbal thought is a topic for both philosophical and empirical study. Pylyshyn (1973), in a metascientific role, has argued most convincingly that at some sufficiently atomic level, the components of imaginal and verbal memory representations must be identical. Simon (1972) has made related arguments. On the other hand, Baddeley and Warrington (1973) have presented evidence showing that long-term memory amnesics lose the capacity to profit from imagery mediation but can still profit from verbal mediation (e.g., taxonomic clustering). Of the three primary researchers of the influence of imagery processes on memory, Paivio (1971) has remained convinced that verbal and imaginal mediation involve separate processing and storage systems, Bower (1972) has converted to a more monistic position (cf. J. Anderson & Bower, 1973), and Rohwer (in press) has decided that research on the other problems of applications for and the developmental study of mediation ("imagery") ability can be pursued independent of a final answer to this first question. We will, following Rohwer, defer consideration of the merits of the dual-process argument until later (see III. B. below).

3. Attention and Foregrounding

There is a third possible explanation for the effectiveness of mental imagery in prose learning. The images that our subjects have in mind as they read may be important as more than a direct mnemonic.

They may function as a foreground (Chafe, 1972) or context for comprehending what comes next. Specifically, we argue that each sentence is understood and remembered relative to a context established by preceding sentences. It is only in this way that the semantic structure of text can be used to facilitate the organization of the text content in memory.

It might seem at first glance that most of the information a person needs to understand and remember a sentence is stored within the sentence. There is presently general agreement that a sentence can be represented as a presupposition and an assertion. The presupposition specifies the context to which the assertion applies. The only other context left to be remembered from sentence to sentence would appear to be the specific objects and people that have been introduced and a general picture of what is going on. That general picture, though, may be very important. Part of understanding a passage involves interrelating all information about a single person or thing. It is necessary to tie each new assertion to previous information about the same referents. Thus, at least portions of that previous information must be kept easily available (in working memory).

One could argue that the task of keeping track of anaphoric (referential) structure is very simple. The characters and objects in a story are highly overlearned early in the course of reading that story. Thus, no one ever makes the mistake of thinking that two references to a particular character are actually references to two different characters. There are, however, limits to our ability to adequately interconnect anaphorically related assertions. Also, the amount of information processing required is often substantial. Our colleague, Perfetti (personal communication), has pointed out that the referents for some nouns used in prose are very diffuse. For example, after a lengthy description of college pranks, one might find the sentence, "The judge frowned on these shenanigans." What highly overlearned verbal response is the referent of shenanigans? Perfetti's example suggests that these problems of discourse reference require generation and maintenance of a general "image" or contextual focus within which the meaning of segments of prose can be represented.

Much recent work has centered on the manner in which subjects generate contexts during sentence and passage learning. Basically, the literature indicates that small units of text are always understood (processed) in a very constructive way. A short paragraph about a concrete event will be comprehended as if it were a picture rather than as a network of the explicit assertions of that paragraph.

Bransford, Barclay, and Franks (1972) have shown this rather nicely. When a sentence such as (1) is presented, it is stored in a relatively concrete form. Thus, even though (1) does not expressly say that the fish swam beneath the log, subjects who have heard (1) know that implicit assertion almost as surely as they know anything else that was explicitly contained in (1). They claim to have seen (2) about as often whether they have actually seen (1) or (3), showing that they have a concrete-semantic rather than lexical-symbolic representation of the sentences they see.

(1) Three turtles sat on a floating log, and a fish swam beneath them.

(2) The fish swam beneath the log.

(3) Three turtles sat on a floating log, and a fish swam beneath it.

From experiments such as this one, it seems certain that context plays a major role in understanding. Sentences are understood when a meaningful representation can be generated for them. Without context, meaningful representation is impossible in many cases. Using similar strategies involving measures of recognition memory, Bransford et al. went on to show that the implications derivable from a constructed semantic representation of one sentence influence our processing of subsequent sentences. For example, a person seeing Passage 4 is about as likely to recognize (5) as a sentence he has seen as to "recognize" (6), even though only (5) was actually seen. This suggests that people integrate the consecutive sentences of sentences like (4) into a single "scene."

(4) There is a tree with a box beside it, and a chair is on top of the box. The box is to the right of the tree. The tree is green and extremely tall.

(5) The box is to the right of the tree.

(6) The tree is to the left of the chair.

Not only does context play a role in prose processing, but it is often essential. For example, Dooling and Lachman (1971) have shown that vague, metaphorical passages are not well recalled unless they are given titles that establish a passage theme or general processing context. Further, the presence of a theme during initial reading, rather than afterwards or on a recall test, is the critical factor in producing better recall (which we interpret as more organized comprehension) of the otherwise disorganized material (Dooling & Mullet, 1973), showing that context plays its role during initial reading.

While we are not convinced that a simple associative structure captures all of the features of human memory representation for comprehended prose, it seems reasonable to think of prose memories as consisting, in part, of associations between context images or themes and the representation constructed for clauses comprehended within those contexts. Thus, each clause, as it is comprehended, becomes associated to the context in which it was comprehended. By context we mean the currently activated portions of the subject's memory, more or less akin to Chafe's (1972) foreground.

All of this suggests that recall of prose depends upon having a theme or context or foreground in mind as one reads each bit of text. If the context is updated constantly, then a retrieval plan is implicit in the memory for that prose. Each context image cues one or more propositions. When these propositions are combined with the old context image, they become the context image for the next propositions, and the process recycles. This view gains some support from the results of Perfetti and Goldman (1974) which show that of the nouns in a sentence, that which is the linguistic theme for the sentence will be the best recall probe.

Such a model has two basic problems. First, subjects (at least adults) seem to show the appropriate behaviors even without extraordinary imagery-related instructions. Data reported by Bransford et al. (1972) and other researchers show evidence of close-to-optimal attention to context or foreground. Second, these studies show no developmental trend; the Bransford et al. design has been extended to seven- to eleven-year-old children with the same basic results (Paris, 1973). Thus, extant data appear to show that even seven-year-old children already comprehend text relative to its thematic or foregrounded information, while our results suggest that this is a skill that must be learned, probably after age seven or so.

One way to solve this dilemma lies in the realization that real reading tasks tend to put larger retrieval demands on the reader. The studies cited above dealt with automatic effects of local context, i.e., information given very near in time to, and in connection with, the same image or episode as the clause currently being comprehended. Retrieval of an entire passage may require a higher level of organization or more ties of the passage information to more stable portions of the subjects' long-term memory.

Indeed, there are both logical and empirical arguments that while understanding and remembering of prose is inferentially complete at the local level, it is not complete globally. First of all, the fact that we can be creative and make discoveries about existing empirical knowledge indicates that there are many inferences which follow from our current knowledge that we do not automatically know. When the volume of information that one acquires is massive, it is only at the local level that comprehension is inferentially complete. The inferences that follow from pieces of knowledge acquired at widely different times are not automatically made. The work of Frase (1969) bears this out. He found that recall of a prose passage contains very few of the inferences that follow immediately from the passage assertions.

One is left to conclude that comprehension of a given text assertion is constructive only with respect to other information that is available (i.e., in working memory) when that text assertion is encountered. If that is the case, then the content of working memory should be a key factor in influencing the structural connections present in memory for prose. J. Anderson and Bower (1973, p. 314) have argued that the imagery effect in sentence learning is due to formation of a larger number of the possible relational associations among the components of the presented sentences. Their experimental data show that the memory structure for sentences learned via imagery instructions is not qualitatively different from that for sentences learned otherwise; it is merely more complete.

We remain neutral on the question of whether there are qualitative differences in memory representations for passages learned with versus without imagery instructions. In either event, the work on context effects and the Anderson and Bower work both lead to the prediction that imagery instructions should produce improved prose learning via changes in the operations of working memory and long-term storage processes. The question of the exact source of the effect is discussed in Section III.

4. Imagery as an Instructional Metaphor

The previous section presented arguments about the possible role of imagery in prose learning that were totally independent of the essence of imagery. That is, they did not invoke any processes that made special use of the sensory-perceptual mechanisms. The arguments spoke of context, of relational connections in memory, of attentional processes, and of short-term memory content. The whole argument could have been made about other instructions, perhaps, that do not even mention imagery. We accept the possibility that imagery instructions may be explainable by theories that ignore the distinction between visual-spatial and verbal-symbolic representation.

Specifically, imagery instructions may be a metaphorical way of teaching certain cognitive processes that we cannot specify in more analytic terms. Even if this is the case, we still must first demonstrate the imagery effect. Then, however, we should direct our attention to understanding the processes that the reader invokes in response to imagery instructions. After we understand the processes involved in prose learning, we may still wish to use imagery instructions, especially with children. However, we would be doing so with some understanding of their effect and some chance of eventually giving the reader more complete self-knowledge of his reading skills.

C. Developmental Arguments

1. The Work of Rohwer and of Levin

We consider now the reasons for expecting a developmental trend in imagery mediation of prose learning. To begin with, we know that imagery mediation abilities show strong developmental trends over the period from age four or five to advanced adolescence. These trends have been discussed at length by Rohwer (in press) and by Levin (1974). Somewhere between kindergarten and first grade, children acquire the ability to use imagery instructions effectively without receiving any motor training (Varley, Levin, Severson, & Wolff, 1973). That is, children in this age group showed improved associative learning in response to imagery instructions; they will benefit from imagery practice even if they do not have the actual objects to manipulate or the chance to make drawings. At the other extreme, Rohwer has shown that it is not until somewhere between the ages of eight and fifteen years that children will spontaneously elaborate during paired associate learning.

We have, of course, argued that imagery does not play the same meaning-enhancing role in prose learning that it has been thought to play in paired associate learning. At the same time, though, the mere fact that there is an important developmental trend in imagery mediation of paired associate learning suggests that there might be a developmental trend in imagery mediation of prose learning. If such a trend exists, one might expect that imagery mediation of prose learning in very young children (say, ages five to six) is impossible; that a certain amount of training is required to bring it out in children in the intermediate ages; and that adults will spontaneously, or with very little instruction, use imagery to mediate their comprehension of prose. Such a finding would strengthen arguments that the same processes are involved in both paired associate and prose learning.

2. Reproductive versus Anticipatory Imagery

A more important argument for the potential existence of a developmental trend in use of imagery to mediate prose learning is suggested by the work of Piaget and Inhelder (1971, 1973). They have demonstrated that many varieties of mental imagery and corresponding memories for those images depend upon the attainment of operational thinking. Specifically, they distinguish between reproductive and anticipatory images. Reproductive images are representations of something that has directly been seen or experienced. Anticipatory images, on the other hand, are "prior imaging of as yet unrealized process" (Piaget & Inhelder, 1971, p. 352). Only the reproductive images are possible before operational thinking has been attained.

Unfortunately, it is the anticipatory images that are most likely to be the kind of images used in imaginal mediation of prose learning. The reason for this is very simple: a reproductive image is an image of some specific event that the subject has already experienced. Prose, in general, is not a record of the reader's experience but rather a record of some experiences of the speaker or the writer that the reader is supposed to try to comprehend. If the reader is to get the actual meaning intended from a piece of prose, he must start with his experience and transform that experience into the event suggested by the text. Piaget and Inhelder suggest that this "anticipation" or generation is exactly what preoperational children cannot do.

The preoperational child, not being able to "anticipate" the meaning intended by the prose, may even be inhibited in his performance by imagery instructions. Such a child may take a verbal message which he would be able to store verbally without difficulty and translate it into a sequence of reproductive images. In such a process of concretizing and individualizing the information, the text's meaning may be distorted. Thus, when asked to recall what it is that he has read or heard, a preoperational child may well end up recalling images that he has experienced but which are not what the prose was referring to. In the developmental literature, this process is often called "over-elaboration."

We can, then, predict a rough developmental trend in the ability to use mental imagery in prose learning. Preoperational children should not be benefited by imagery instructions in the learning of large bodies of text; indeed, they may do worse trying to have images. Children who have acquired the concrete operations should become trainable in the use of mental imagery to mediate their comprehension. Finally, adults, as practiced readers, should automatically, or very close to automatically, be making use of this tool.

The experimental results reported herein bear out these predictions. However, it is interesting to note that the negative prediction for preoperational children's use of imagery may be counterintuitive. One colleague, for example, reports that he still remembers a picture in a first-grade book showing how "the mouse ran up the clock." He argues that all of the pictures, printed and mental, that appear to children reading storybooks must facilitate comprehension. They do, we suppose, if the material is simple enough. The child probably does have "mental pictures" corresponding to storybook illustrations and past experiences cued by the text.

If the pictures are given, or if the text is meant to convey information about specific episodes the child has already encountered, or is encountering in a single, specific picture, then there should be a positive effect of imagery. On the other hand, the information being conveyed may go beyond single experiences of the child. If so, he will not be able to construct (anticipate) a correct representation. Even worse, the text may contradict some portion of the child's related experience. In this case, the child's image should interfere.

II. Preliminary Empirical Studies

A. General Format

The five experiments reported here are preliminary studies, designed primarily to demonstrate the presence or absence of imagery instruction effects in children of different ages and in adults. As such, they are not uniform in design. There are, however, certain concerns that resurface throughout the studies, and we will consider them briefly before presenting the studies individually.

1. Basic Procedure and Scoring

All of the experiments reported in this paper share a basic design aspect. Specifically, subjects read or hear prose passages in different experimental situations. Then they are asked to recall, without regard to exact wording, all of the information that they can remember from the passages. The measure of recall is based on the number of text propositions recalled out of the number possible. Since this scoring procedure is common to all five experiments, we will describe it separately.

The passages which our subjects read and recall are analyzed in advance for their propositional content (see page 19), and a checklist is prepared that lists all of the separate propositions of the text. By the time such a list is prepared, two or three members of the research team have reached agreement on the completeness and accuracy of the list. The list is then used as a scoresheet for each subject's recall protocol. One or two judges check off which propositions are represented in both the list and the particular recall protocol. The recall score is taken as the percentage of propositions recalled out of the total in the list. The procedure seems to be reliable and valid.

Validity can be considered at two levels. First of all, is a reconstruction task a good measure of comprehension? We think that it is. What could be a better measure of the acquisition of knowledge than the ability to reconstruct that knowledge? Certainly there are comprehension tasks that put a low load on memory, such as instruction following. However, most intellectual activity involves the synthesis of knowledge acquired from various sources in the past. Thus, much reading is done for the purpose of acquiring relatively long-term use of information. Certainly this is the case with much school-related reading. Therefore, the ability to demonstrate reconstructive knowledge of passage content seems to be an obvious criterion for adequate acquisition of prose.

At the more specific level, given acceptance of a memory test for comprehension, we must ascertain whether the scoring procedure we use is valid. The procedure's validity rests on the method we use for deriving propositions from prose passages. We look for four classes of propositions: verbal predications, adjectival predications, class inclusions (nominal predications), and interpropositional relations.

A verbal predication is determined by each noncopula verb in prose. There are both one- and two-place verbal predications, corresponding roughly to intransitive and transitive verbs. Either sort is counted as one proposition. There are also two types of adjectival propositions that we count, attributive adjectives (e.g., "John is ingenuous") and adnominal adjectives (e.g., "The tall man . . ."). Each specification of a superordinate categorization (e.g., "John is a barber") is counted as a proposition. Finally, relations between propositions (often indicated by subordinate clause forms) are themselves scored as separate propositions.

For example, given the sentence, "Mary is a star quarterback because she eats Wheaties," we would specify the following four propositions:

(a) Mary is a quarterback
(b) star (quarterback)
(c) Mary eats Wheaties
(d) (a) because (c)

If this sentence appeared in a criterion passage for one of our experiments, judges would score a recall protocol by checking off which of the four propositions actually appeared in the protocol. The final measure used is the percentage of propositions recalled out of the total number on the checklist.

Even though a number of problems arise in scoring, inter-observer reliability about the presence or absence of single propositions is over 80 percent. The only data reported here are total scores. For these, inter-observer reliability is about 97 percent. Consequently, we usually have a single scorer, checking random samples of the scoring with a second scorer. Thus, even though there may be great subjectivity in deciding what propositions were recalled (when, for example, a subject sees "Mary is a star quarterback" and recalls "Mary is a star"), the total scores are reliable.

2. Appropriate Controls

The appropriate control groups for experiments such as those reported here are hard to specify. If, as in Experiment 1, the imagery induction is only an instruction, an appropriate control is a condition that is the same except for lack of the imagery instruction. Our problems arise because of our wish, in Experiments 2, 3, and 5, to more strongly assure imaginal processing. In those experiments, we forced the imagery processing by requiring an overt pictorial production. Here two control problems arise. First, nonimagery controls should still impose some task beyond the reading of passages. Second, it is necessary to determine whether any imagery effects are due to processes during reading or to processes during the pictorial requirement.

The first problem is met rather easily. Instead of drawing or constructing pictures relevant to the story they have read, control subjects do other drawing tasks that (a) fill the time interval, (b) involve drawing, (c) are unrelated to the criterion passage, and (d) have no verbal-semantic content that could interfere with the passage content. Such tasks include geometric form drawing, solving of printed mazes, etc. In Experiment 2 we also used a control group that alphabetized the words of the passage instead of drawing, in order to rule out the possibility that doing drawings per se may have an interference or a motivational effect.

The other control problem is not faced in the present studies. This is the question of whether our imagery effects are the result of processing during reading or during the drawing task. In experiments done since the ones reported here we have attacked the problem; preliminary analyses indicate that if subjects expect the drawing task but do not actually have to do it, they will perform about the same as if they had done the task. These results are incomplete and will be reported in later papers. For now, the problem of determining the locus of effects demonstrated below must remain.

B. Adult Studies

We attempt to answer the first question posed at the beginning of this report: Is there an effect of mental imagery on prose learning? This initial demonstration was performed on adult subjects since they are (a) less variable in their behavior in the laboratory, and (b) an appropriate source for information about the mature reader.

1. Experiment 1

a. Method

Design and materials. Two types of study booklets were prepared. Both had instructions on the first page to read the following story carefully as a recall test would follow. The second and third pages contained Bartlett's (1932) "War of the Ghosts" story. Subjects in the control group (N=17) received exactly this booklet. The experimental group subjects (N=16) received booklets that also included on page 1 the information that keeping in mind a mental picture or image of what was going on in the story would help them learn it better and, hence, do better on the recall test. The design is summarized in Figure 1.

Subjects. Thirty-three subjects participated in this experiment as a voluntary activity of a partially home-study introductory psychology course. Their average age was greater than for the usual introductory psychology class: Median = 21, Mean = 27. They were run as a group.

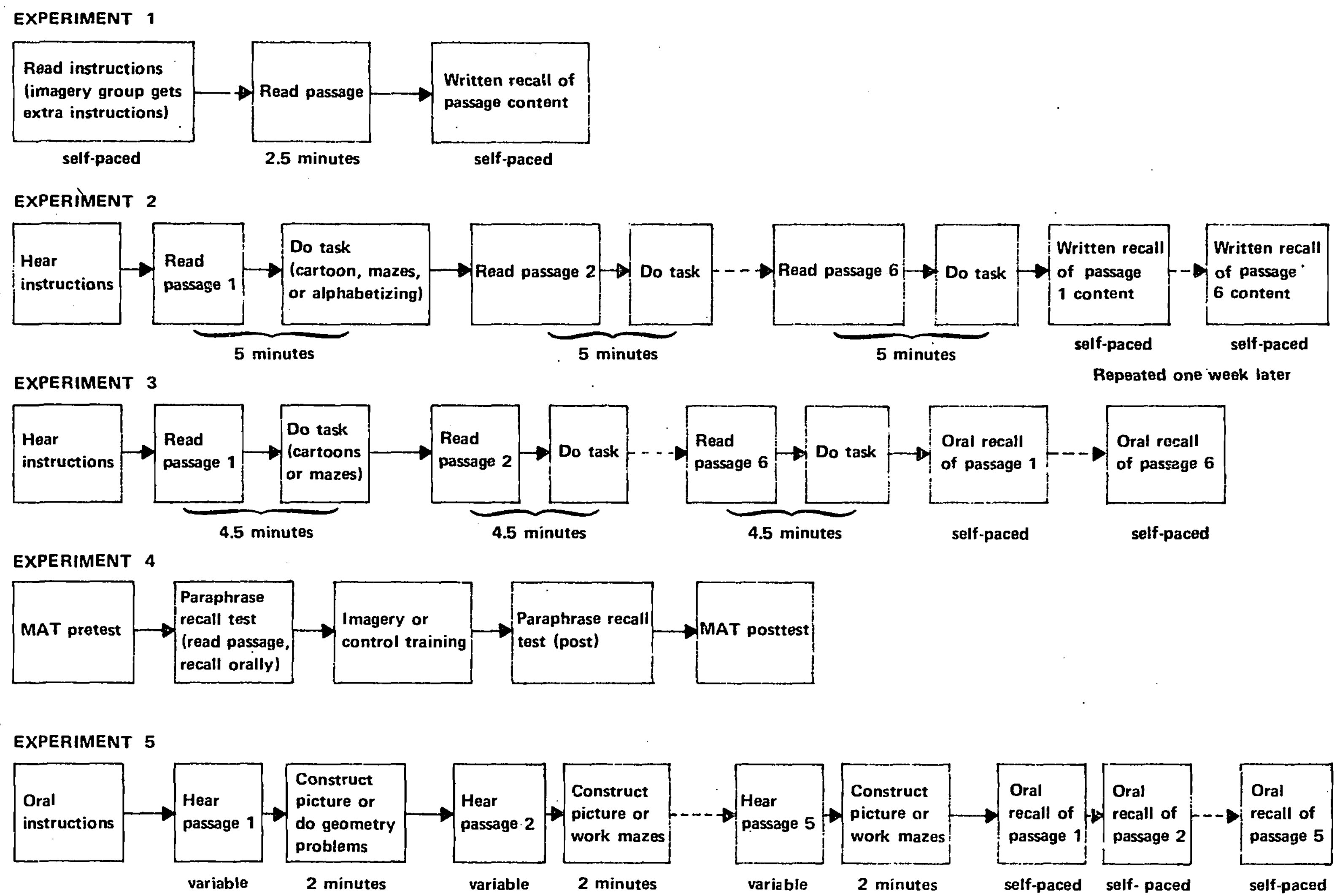

Figure 1. Flowcharts of Experiments

Procedure. The subjects were seated in a modern, well-lit classroom. Alternately seated participants received experimental materials and the remaining people received control materials. The subjects first read the instructions on the front page of their booklets. They were then given 2.5 minutes to read the Bartlett passage. When a stop signal was given, subjects were told to mark how far they had gotten, though many of them forgot to do this. Subjects were then given a blank sheet of paper and instructed to write down all of the information, main facts and details, that they could recall, using their own words as desired.

b. Results and Discussion

The main results are shown in Table 1. Basically, total performance was equivalent in the two groups, $F < 1$. Further, performance was generally better in recall of the first half of the story, $F(1, 30) = 58.7$, $p < .001$. Finally, there was an interaction of story segment with condition such that the experimental group did better on the first half but worse on the second half of the story, $F(1, 30) = 4.26$, $p < .05$.

TABLE 1

Mean Percentage of Propositions Recalled (Experiment 1)

	Imagery Instructions (N=15)	Control Instructions (N=17)	Combined
First half of story	44	37	40
Second half of story	13	18	16
Total	28	28	28

Almost all of the subjects remarked that they had not been given enough time to read the whole passage. This, and the fact that imagery group performance was higher in only the early part of the passage, suggests that the imagery group read at a slower rate. Consequently, even though their recall was higher for the portion that they read, their total performance matched that of the control group. Within this experiment, then, both groups learn the same number of propositions per unit time interval.

If imagery-using subjects always learn and retain the same number of facts per unit time interval as controls, then there seems little reason to be interested in imagery procedures as a means of improving comprehension and memory. However, this may not be the case. Imagery mediation takes time. When study time is drastically limited in paired associate experiments, imagery effects disappear or are attenuated (Paivio, 1971, pp. 343-344). We speculate that in prose learning experiments as well, when time is drastically limited, what is gained by imagery is lost by the decreased amount of text read at all.

The next experiment was designed to look for an imagery-mediation effect in a relatively self-paced study situation. A positive finding here would suggest that once a study time threshold is exceeded, imagery mediation does result in superior prose learning compared to the unspecified strategies of control groups.

2. Experiment 2

In this second experiment, an attempt was made to give self-paced reading time and to manipulate several other variables that seemed relevant to investigations of imagery and verbal behavior. In addition, the imagery manipulation was strengthened and more attention was given to the problem of providing adequate control conditions.

Instead of just telling subjects to make images in their heads, we made the demands much more concrete. They were told that after reading the passage at their own rate they were to spend the rest of a five-minute interval drawing a cartoon strip that conveyed the information of the passage. This was to be done on a sheet with several blank squares that was similar to a comic strip whose picture squares are empty. Subjects were told that keeping a picture in mind of what is happening in the passage as they read it would help them both on the cartoon task and on later recall. Refer to Figure 1 (p. 22) for an outline of the experimental design.

Providing control groups is a serious problem in this research program. One certainly wants to control the time parameters of the experiment, so that time from reading to testing and from passage to passage is constant over conditions. This was done by using a group that was given no imagery instructions and that spent their interpassage time doing nonverbal drawing tasks such as solving pencil mazes and copying geometric figures. Thus, while the imagery group was drawing a meaningful cartoon, the drawing control was drawing material that was semantically empty.

The problem with the drawing control is that it does not control the total amount of time that a subject spends thinking about a particular passage. One could argue that the drawing control group has not been thinking about the passages while drawing the mazes, whereas the imagery group must continue thinking about the passage as they cartoon it. This suggests that another appropriate control is one in which there is no cartooning or imaging demanded but in which the passage remains available and receives some of their attention. We tried to achieve this by a condition in which subjects read each passage ad lib and then spent the remaining interpassage time with the passage in front of them, alphabetizing its words.

Two factors were systematically varied in this experiment: passage concreteness and retention interval. Imagery has been associated with noun concreteness in the verbal learning literature. Thus, there are a number of explanations of the imagery result in prose learning that would call for an interaction of imagery instructions with passage concreteness. Dual-process models would predict that imagery only helps with concrete materials. Theories that emphasize relation of to-be-learned experience to one's concrete, experiential base might suggest that concrete passages cannot be further helped but that abstract passages might benefit from a search for a concrete representation. Thus, the variable was worth including in hopes that some otherwise plausible explanations of imagery effects in prose learning can be confirmed or rejected. Three concrete and three abstract passages were used, the distinction being based upon a count of concrete and abstract referents in the passage.

The factor of retention interval was also manipulated within subjects. This was done to assure that any imagery effects that were found could be examined to see if they are long-lasting. There is very slight evidence (Olton, 1969) that while verbal mnemonics improve noun-noun paired associate learning, slightly more forgetting occurs when material is learned with the sentence mnemonic than when it is learned by rote to a similar criterion. In the present experiment, both immediate and one-week-later recall for passage content was tested.

a. Method

Design and materials. Six criterion passages of 110-116 words each were constructed for use in all conditions. The passages contained 37 to 53 propositions (see II.A.1) each, Mean = 41. Three of the passages were relatively abstract; viz., they contained almost no concrete references. The topics of these passages were the philosophy of Immanuel Kant, the nature of intelligence tests, and reasons for studying history. The other three were concrete; viz., they dealt with specific events, places, and names. These passages were about a military incident, a tragic accident, and a picnic. The six passages were assembled into six different orders, according to a Latin Square design in which concrete and abstract passages were alternated.

In addition to the within-subject variable of concreteness, subjects were also tested immediately and after one week. Finally, there were three conditions: imagery instructions and cartoon drawing, a maze-drawing control, and an alphabetizing control, as described below.

Subjects. Thirty-six subjects participated in this experiment in partial fulfillment of a laboratory requirement for an introductory psychology course. They were assigned alternately to the three conditions according to the order of their arrival at the laboratory.

Procedure. Each subject was given general information about the experiment and instructed to read the passages carefully because he would later be required to write down the information from each of them. The subject then went through a cycle for each of the six passages in which he first read the passage at his own rate and then spent the remainder of a five-minute interval doing one of three tasks (see Figure 1 [p. 22] for schematic description of the procedure).

Subjects in the cartoon drawing condition were to illustrate each passage's content in a series of cartoon pictures that were drawn during the remainder of the reading period on a separate sheet without further access to the passage. Subjects were told to keep in mind what they would draw as they read the passage.

Subjects in the *maze* drawing condition spent the remainder of their five-minute reading/activity intervals doing a series of paper and pencil mazes and copying complex geometric figures, all without further access to the passage just read.

In contrast to those in the other two conditions, subjects in the *alpha*betizing condition kept the passages in front of themselves during the whole five minutes. After reading the passage ad lib, these subjects turned to the next page of their booklets, where they found a second copy of the passage plus a form with word-size blanks. Their task was to write the words of the passage into the blanks in alphabetical order.

After completing the reading and tasks for all six passages, subjects in all conditions were asked to write down in their own words all of the information they could remember from the passages. They were told not to worry about syntax, punctuation, or neatness, but to write down every fact, including adjectival information. Each passage was prompted in turn for this unpaced written recall by presenting a two-word cue for the passage.

Subjects returned one week later to repeat the recall test.

b. *Results*

The means of the various treatment combinations for which significant ($p < .05$) differences were found are given in Table 2, along with the appropriate F-statistics. The main effect of the cartoon condition was significant ($p < .05$), confirming our expectation that with longer study time, a strong imagery condition would produce better performance than nonimagery controls. There are both retention and story effects ($p < .01$), as one might expect; concrete passages are better remembered than abstract passages and immediate recall is better than delayed recall.

TABLE 2

Mean Recall Percentage for Significant Treatment Effects (Experiment 2)

Effect	F-Statistic	Means	
Conditions	$F(2,42) = 3.38^{*}$	Cartoon: 35	
		Maze: 26	
		Alpha: 29	
Retention Interval	$F(1,474) = 35.5^{**}$	Immediate: 33	
		Delayed: 27	
Stories	$F(5,74) = 77.8^{**}$	Abstract	
Abstract vs. Concrete:	$F(1,74) = 181^{**}$	Kant: 19	
Residual:	$F(4,74) = 52.0^{**}$	IQ: 26	24
		History: 27	
		Concrete	
		Army: 25	
		Accident: 39	36
		Picnic: 42	

	Cartoon	Maze	Alpha	
Abstract				
Kant:	23	16	20	
IQ:	30	25	25	
History:	35	23	24	Conditions by Stories $F(10,474) = 2.36^{**}$
Concrete				
Army:	31	22	22	
Accident:	46	36	37	
Picnic:	46	36	46	

* $p < .05$

** $p < .01$

There was no interaction between conditions and retention interval, $F < 1$. In fact, the percentage of initial recall retained one week later was almost identical for the cartoon, maze, and alpha conditions: 84.7 percent, 82.9 percent, and 84.1 percent, respectively. The cartoon effect, then, appears to lie in better initial acquisition of information. If there are retention differences, they must be in very short-term (30 minutes or less) retention.

The final entries in Table 2 show the significant ($p < .01$) interaction between conditions and stories. This results from the inordinately high performance of alpha subjects on the Kant and Picnic passages. Such a finding, presently inexplicable, suggests that we should properly treat passages (stories) as a random treatment factor, since the six stories we have used are only a small sample from the set of all possible passages. When this is done, there is no simple F-ratio that is appropriate for testing the conditions effect. Rather, a quasi-F-ratio must be used (Clark, 1973). This statistic shows that the probability of the conditions effect being replicable with new stories and new subjects is not as high, $F'(2, 52) = 2.88$, $p < .10$, as the above-reported probability that replication would occur with new subjects but the same stories, $F(2, 42) = 3.38$, $p < .05$.

c. Discussion

The present results appear to confirm our expectations that imagery instructions facilitate the learning of information from prose. The effect is isolated as one of acquisition or short-term retention but not long-term retention, as there is no interaction of conditions with retention interval. In addition, both concrete and abstract passages are facilitated. An examination of the data reported in Table 2 shows that the amount of facilitation of the cartooning task over the other two tasks is about the same for abstract as for concrete stories. The specific figures (in percents) are:

	Cartoon	Controls (Pooled)	Differences
Concrete	41	33	8
Abstract	29	22	7
Differences	12	11	

There is a large concreteness effect and a cartooning effect almost as big, but no interaction.

This suggests that the effects of imagery-cartoon instructions do not lie in their making available a second nonverbal coding system, as Paivio (1971) has proposed for paired associate learning. If that were the case, we would expect an imagery effect mainly for the concrete passages, since they are the ones most amenable to representation as images rather than as symbol strings. However, there is no such interaction, suggesting that the dual-coding explanation is insufficient. In fact, there is no reason to believe that we would not get the same effect in adults by requiring verbal summaries and telling subjects to keep in mind what they will say in their verbal summaries as they read.

Even if there should be no advantage in adults of nonverbal elaboration instructions over verbal elaboration instructions, it is still useful to know that elaboration procedures help. Requesting subjects to have a picture in their minds as they read and to be ready to draw a cartoon may be more certain a procedure than requesting the equivalent verbal performance instructions. Certainly, we can expect this to be so with younger subjects. The picture-in-your-mind metaphor may tell the subject what to do better than a "verbal" instruction, regardless of what the critical performance is. Further research is underway in this project to see whether any effects of imagery versus verbal summary can be detected, either in the particular types of propositions recalled or in interactions of the verbal-nonverbal factor with passage structure or content.

Because of the order of events in this experiment, an alternative explanation is possible. This alternative states that the task demands imposed in the three conditions permit differing amounts of time to be spent thinking about passage content prior to the recall test. Thus, subjects in the cartooning condition could be learning during the cartoon-drawing time as well as during the reading time. We are now conducting experiments designed to discover whether or not learning actually takes place during cartoon-drawing time. In these experiments, either the order of the cartoon and recall tasks is reversed, or the expected cartoon or verbal summary task is replaced at the last minute by an alphabetizing task. We have not finished collecting data; but the available data do not support this time-on-task explanation.

C. Mental Imagery and Comprehension in Third and Fourth Graders: Experiment 3

The preliminary hypotheses we have been developing for the comprehension process (see I. B. 3) emphasize that a good deal of attentional and organizational effort must go into the mature comprehender's reading under imagery instructions. While reading, the student must keep in mind (i. e., in working memory) an organized representation of what has been comprehended so far. At the same time, he processes the phrase or sentence currently being read. Further, the student must know enough about the spatial arrangement of the page so as not to lose his place. Because of his tendency toward centration (Ginsburg & Opper, 1969, p. 167), the child of age seven or less may not be able to attend simultaneously to these three dimensions of local processing, page spatial organization, and maintenance of a complete and organized foreground. As we shall discuss in Section II. E., the fact that the student is required to attempt to work simultaneously in these three domains may hinder his ability to perform well in any of them.

Though they may not have been designed with exactly these thoughts in mind, primary-grades reading series often recognize the six-year-old's inability to multiprocess by using very short text units--single phrases to short paragraphs. In this manner, no previous context is present (other than overlearned familiarity with a few recurrent characters) and there is very little spatial organization required. If the spatial organization of even this small amount of text is difficult for the child, he will often use a finger to mark his place. A new reading instructional system under development at LRDC intentionally uses pictures to similarly externalize comprehension foreground or context (Beck, personal communication).

Previous work has usually suggested that pictures do not help, and may even hinder, comprehension (Samuels, 1970). However, the specification of the procedure for generating pictures and their relation to the text has not been systematically controlled. For some authors, pictures are motivators, for some they are an alternative source of information; it is also possible that pictures can function as external memories for foreground information. In future research, this group plans to test the use of pictures when they are designed to function as external context (foreground) memories for children trained to use them as such.

The present experiment is merely a general examination of whether children who, in theory, have just gained the ability to maintain the several aspects of cognitive activity that reading requires are able to benefit from imagery instruction.

1. Method

Design and materials. Each subject read the same six passages, though the order of passage was counterbalanced via a Latin

Square design. Each passage was of approximately 100 words in length, of easy readability for third graders, and entirely about concrete objects and specific events. There were approximately 36 predications per passage, according to the scoring scheme discussed in Section II. A. 1. Two conditions were run, corresponding to the cartoon and maze conditions of the previous experiment.

Subjects. Twenty third-grade students from a campus laboratory school and 20 fourth-grade students from an urban parochial school participated, half in each condition. Subjects were assigned to each condition on the basis of standardized reading scores, such that the mean for the experimental group was the same as the mean for the control group.

Procedure. Subjects in each condition were told to read the six passages very carefully since they would later be required to tell what the passages were about. Subjects were then given instructions concerning the nature of the task to be performed after reading. A total of 4.5 minutes was spent on reading and working on the task for each passage. See Figure 1 (p. 22) for a flowchart for this experiment.

The cartoon subjects read each passage at their own rate, and then drew pictures illustrating the content of the passages for the remaining time. Subjects were instructed at the beginning to keep their drawings as simple as possible. They were not allowed to refer to the passage while drawing. Drawings were made on separate sheets which had been divided into four blocks to resemble a newspaper comic strip.

The maze subjects also read the passages at their own rate, and spent the remaining part of the 4.5 minutes working on a series of nonverbal drawing tasks, such as solving printed mazes and copying geometric figures.

Prior to the third passage, all subjects were reminded to read carefully because they would later be asked what the passages were about.

Following the presentation of the six reading-drawing cycles, the subjects were given a key word from each passage as a prompt, and were asked to recall what the passage was about. Recall was recorded on tape. Subjects were run individually.

2. Results and Discussion

Recall from each story was expressed as a percentage of the total number of propositions possible in that story. A second score was derived by dividing the number of picturable propositions recalled by the total number of picturable propositions in each passage. Separate analyses of variance were performed for each measure, with independent variables of subject's school and experimental condition between subjects and passages within subjects.

The analysis of total predication recall yielded no significant main or interaction effects ($p > .25$), except for a passages effect, $\underline{F}$ $(5, 180) = 16.26$, $\underline{p} < .001$. In particular, the experimental group performance of 36 percent was not significantly better than the control group's 31 percent.

The analysis of picturable proposition recall yielded some significant differences. None of the between-subject effects were significant, $\underline{F}$ $(1, 36) \leq 1.55$, $\underline{p} > .05$. However, there were several significant within-subject effects. Passages was highly significant, $\underline{F}$ $(5, 180) = 40.7$, $\underline{p} < .001$, as were the interactions of passages with schools, $\underline{F}$ $(5, 180) = 4.91$, $\underline{p} < .01$, and with conditions, $\underline{F}$ $(5, 180) = 2.36$, $\underline{p} < .05$. The three-way interaction was not significant, $\underline{F}$ $(5, 180) = 1.93$, $\underline{p} < .10$.

The means for these effects are shown in Table 3. The interaction effects appear to be the result of school differences on passages A, C, and D, and condition differences on passages A, C, and F.

Reading time was also recorded for each subject on each passage. As has occurred often (but not always) when we have measured reading times, the times were slightly greater for the experimental than for the control condition (Means = 56.4 seconds, 47.3 seconds), though the effect was not significant, $\underline{F}$ (1,36) = 2.71, $\underline{p} > .10$. All other effects were even less likely to be reliable, $\underline{F}$-statistics < 1.12. Further, there was no correlation between reading time and performance, both measures being totalled over stories for each subject, $\underline{r} = .10$, $\underline{N} = 40$, $\underline{p} > .20$.

TABLE 3

Mean Percentage of Picturable Propositions Recalled

	Passages					
	A	B	C	D	E	F
Urban Parochial 4th Graders						
Experimental	44	48	68	36	39	41
Control	37	47	57	38	32	43
Total	40	48	62	37	36	42
Campus Lab 3rd Graders						
Experimental	55	46	60	50	42	49
Control	40	44	54	43	35	40
Total	48	45	57	46	38	44
Total Experimental	50	47	64	43	40	45
Total Control	38	46	56	41	34	42

It is possible that with extremely large samples we would find an advantage for the experimental group over the maze control group. In any event, though, the effect is certainly not the clear, substantial effect seen for adults in Experiment 2. This suggests that either the children lack some prerequisite ability for benefiting from cartoon instructions or they need more familiarization and practice in either the cartooning task or the relationship between that task and their reading process. The next experiment examines the result of a training study that provided this practice.

D. Mental Imagery Training for Third and Fourth Graders: Experiment 4

In this section, we describe a training procedure which takes third- and fourth-grade children from an initial inability to benefit from imagery task instructions to a point at which they do substantially better in reading when given imagery instructions than when not instructed to use imagery. The behavior of the trained children is essentially the same as that of adults.

The training procedure was built around the primary goal of giving the child a clear understanding of what his mental images should consist of as he reads. Thus, the children were given many passages to read; after each one, they were asked to immediately draw a cartoon sequence (a set of stick-figure pictures) that illustrated the passage's content. Later in the program, an attempt was made to transfer the teacher-originated criteria for cartoon adequacy to internal criteria for undrawn mental images.

Over the course of the training program, the criteria for children's cartoons were gradually tightened to make each cartoon a relevant and complete semantic image of its passage. A number of devices were used to do this. Initially, emphasis was placed upon segmentation of the story into picturable events and illustration of each segment in a separate square of a blank cartoon form. Emphasis was placed on simple, quick, stick-figure drawings in order to overcome the children's tendencies to become captivated by details (e.g., heels on shoes, buttons on clothes, etc.).

Once the basic objective of getting each child to produce a reasonably relevant multi-picture cartoon strip was attained, we taught the children to make some important discriminations in the course of their reading and drawing. They were taught to select the main points in each passage for use as main themes in their cartoons. The basic objective was for each cartoon scene to represent a main point and its related details, and for the child to be aware of unpicturable facts (e.g., "John hoped he'd win the contest") that related to the main points.

A series of group and individual exercises followed in which the children were given specific criteria for judging whether their cartooning and imaging were adequate. For example, each sentence in the story needed to have some representation in the child's cartoon. The children took turns showing what parts of their cartoons corresponded to particular sentences of the passages. The eventual goal was for each paragraph to correspond to one scene in the cartoon.

The children were next given practice in using their cartoons to recall the main points, details, and nonpicturable content of the passages. After this, children practiced recalling passage content without having their cartoons in front of them. Finally, on the last day, the actual cartooning was replaced with exercises in which the children were asked to imagine what they were planning to draw as they read (although they were not required to actually draw a cartoon). Also, on the last day recall after reading with imagery was practiced.

All of this procedure is predicated upon the assumption that children in the third and fourth grades have the basic mental capacity for benefiting from the imagery strategy but not the specific learned skills. The training procedure we adopted was intended to teach those skills relatively quickly and efficiently.

1. Method

Design and test materials. Subjects were first pretested with the Metropolitan Achievement Tests Elementary Reading (MAT) Form F and a paraphrase recall test. The MAT was selected because it is commonly used as a measure of general reading ability. The paraphrase recall test consisted of a three-paragraph, 176-word passage which the subject first read and then retold from memory in his own words. The story used simple vocabulary and a low enough syntax level so that the children would be comfortable with the passage. Two forms of the story were used that had equivalent semantic structure but featured different characters (e.g., a rabbit in the carrot patch versus a beetle in the rose bushes). Recall protocols were scored as outlined in Section II. A. 1.

Based upon these test scores, the subjects were split into two groups that were approximately identical in mean MAT and paraphrase recall pretest scores as well as grade level and sex distribution. (Pretest data are given in Table 7.) One group received twelve imagery training sessions while the other group received equivalent reading practice sessions, as described below. On two different days after the last training session, subjects were given three posttests: MAT Form G, paraphrase recall, and paraphrase recall with imagery instructions.

Both paraphrase recall tests had alternate forms of equivalent semantic structure, but the forms for any one of the three paraphrase recall tests were not equalized for difficulty with the forms for either of the other two tests. The first paraphrase recall posttest passages contained 230 words over five paragraphs and were of somewhat greater reading difficulty than either the other paraphrase recall passages or the instructional materials for the training sessions. The second paraphrase recall posttest was given with imagery instructions. The dual passages for it each contained 182 words over five paragraphs.

Subjects. Thirty-two children (10 third graders and 22 fourth graders; 20 girls and 12 boys) from an inner-city parochial school served as subjects. Seventy-five percent of the subjects were below grade level on the Elementary Reading subtest of the Metropolitan Achievement Tests. They were assigned to the two groups according to a scheme that equalized the groups on the basis of sex, grade level, and mean pretest scores.

Procedure and content of instructional materials. The experiment covered a period of 58 days. The schedule of events is summarized in Tables 4 and 5 and in Figure 1 (p. 22). The MATs were given according to the standard instructions. All three paraphrase recall tests were preceded by instructions in which subjects were told to read the story and that they would be asked to tell the experimenter what they remembered from the story. After self-paced reading, usually aloud, the children orally recalled the story content. The second paraphrase recall posttest included additional instructions for subjects to picture in their minds what is happening in the story. Training took place in groups of five or six children. MAT testing was in classroom groups, and recall testing was done individually.

TABLE 4

Summary of Experimental Training Procedure

Day	Number of Stories	Paragraphs per Story	Activities
1	-	-	MAT pretest.
23	-	-	Paraphrase recall pretest.
24	2	1	Subjects read first story and are shown cartoon strip E has drawn. Fast, stick figures emphasized. Teacher points out relevant scenes in sample story and subjects draw their own version of each scene in turn. Emphasis on illustrating story, not subjects' fantasies. Subjects then read and illustrate another story.
28	2	1	Subjects read and illustrate both stories. Experimenter checks pictures for accuracy, detail, and order of events. Emphasis on something in cartoon for each sentence of story.
35	2	2	Same as previous session plus emphasis on what in picture corresponds to each paragraph, and how many scenes story contains. Call on one child to tell story from his picture.
36	2	2	Same as previous session with emphasis on close correspondence between pictures and story. In addition, subjects are told to mark pictures to show points where unpicturable statements occurred in story.
37	1	2	Same as previous day.
38	1	3	Same as previous day.
39	1	5	Same as previous day, but each subject tells story from his pictures. Emphasis on recalling all details; feedback on recall ability.
42	1	3	Summary to subject of skills he has learned. Each child gets a different story. Otherwise, same as previous day.
44	1	3	Same as previous day.
45	1	5	Same as previous day.
46	1	5	All get same story. Each subject tells story from his cartoon. Then cartoons are compared frame by frame.
50	1	5	Subjects told to have images instead of drawing. Each child tells story, then draws his images as cartoon. Subjects told of forthcoming tests and that training sessions will help.
51	-	-	Paraphrase recall posttests.
52	-	-	MAT posttest.

TABLE 5

Summary of Control Training Procedure

Day	Number of Stories	Paragraphs per Story	Activities
1	–	–	MAT Pretest.
2	–	–	Paraphrase recall pretest.
24	2	1	Subjects read first story and are asked multiple-choice questions and shown how to mark answers. Then subjects read another story and answer questions, which are checked.
28	2	1	Subjects read stories and answer questions on paper. Then questions discussed orally. Emphasis on accuracy.
35	2	2	Same as previous day. Subjects told that this procedure will improve their ability to remember what they read, i.e., if you can read well enough to answer questions, you can read well enough to recall whole story.
36	2	2	Same as previous day.
37	2	2	Same as previous day.
38	2	3	Same as previous day.
39	2	5	Same as previous day.
42	1	3	Same as previous day, but each child gets his own story.
44	1	3	Same as previous day.
45	1	5	Same as previous day.
46	1	5	Same as previous day, but one story for all.
50	1	5	Same as before, but separate story for each child. Subjects told of forthcoming tests and told that their training sessions have helped them prepare for it.

The experimental treatment consisted primarily of training in reading stories and then drawing cartoons (sequences of stick-figure sketches) of these stories from memory. Each session lasted about 25 minutes. Details of the instructions are given in Table 4. The control treatment consisted of practice in reading passages and then answering multiple-choice questions about them. Each session covered about 15 minutes. On any training day, control subjects read all of the stories of the experimental group and sometimes (in Sessions 5-7) more. Details are shown in Table 5.

The experimental treatment was specifically designed to provide instruction in the following skills:

(a) Drawing of stick figures.

(b) Segmenting stories into separate, picturable scenes.

(c) Discriminating main events from details.

(d) Discriminating picturable predications from nonpicturable.

(e) Relating each sentence in the story to a portion of a picture.

(f) Relating each paragraph in the story to a portion of the picture series.

(g) Relating each picture to a portion of the story.

(h) Gaining an approximate one-to-one correspondence between paragraphs of the story and pictures in the series.

(i) Using pictures to recall the main events of the story.

(j) Using pictures to recall the details of the story.

(k) Using codes on pictures to cue recall of unpicturable propositions.

(l) Substituting imagery for actual pictures in the above objectives.

The stories used in the training sessions were selected to be easily readable by third-grade children and amenable to illustration in cartoon strips. In particular, the following principles were observed: (a) vocabulary should be easily decodable or clearly part of the sight vocabulary of the third graders; (b) both vocabulary and syntax should approximate oral speech patterns for third graders; (c) the stories should represent a series of events involving picturable characters; and (d) dialogue should be minimized.

2. Results

Means and standard deviations (in parentheses) for the two conditions on all pre- and posttest measures are shown in Table 6. As can be easily seen, the only distinction between the two groups is on the posttest paraphrase recall task with imagery instructions. Analysis of variance and covariance analysis with the two pretests as covariants both show that measure to be the single posttest distinction between the groups. The analyses are shown in Table 7. The relationship between performance during training and posttest performance is still being studied.

TABLE 6

Means on Pre- and Posttests
(Standard Deviations in Parentheses)

Condition	N	Pretests		Posttests		
		MAT	Paraphrase	MAT	Paraphrase	Paraphrase with Imagery Instructions
Control	16	64.2 (11.1)	0.30 (0.12)	65.0 (12.2)	0.21 (0.17)	0.31 (0.19)
Experimental	16	64.2 (8.4)	0.31 (0.09)	65.1 (8.3)	0.25 (0.17)	0.44 (0.11)

TABLE 7

F-Statistics from Analyses of Variance and Covariance (See text)

Test	df	Pretests		Posttests		
		MAT	Paraphrase	MAT	Paraphrase	Paraphrase and Imagery
Analysis of Variance	1, 30	0.00	0.15	0.00	0.53	5.85*
Analysis of Covariance	1, 28	--	--	0.00	0.44	9.42**

* $p < .025$

** $p < .01$

3. Discussion

There was some success with the training procedure in our initial efforts. When the imagery instructions were given as a prompt, the experimental treatment group did 40 percent better than the control (44 percent versus 31 percent). Thus, we can conclude that the experimental procedure trains subjects to do something useful to their comprehension skills, but it does not get them to the point where they invoke that skill without a specific instruction to do so. It would be even more desirable, of course, if the training procedure produced spontaneous use of improved reading comprehension strategies. This did not happen in the present study.

Of course, we should not expect complete spontaneity of these processes, since our own data show that adults do not invoke them spontaneously. Adults (see Section II) appear to benefit from instructions to have an ongoing image of what is going on in a passage as they read it. Thus, they have skills similar to those that we appear to have taught younger children. On the other hand, the mere existence of an instructional effect shows that adults are not completely spontaneous in their use of the strategy corresponding to imagery instructions. If they spontaneously used the strategy, then telling them to do so would not improve their performance.

The fact that there was no change in standardized test (MAT) performance can be interpreted two ways. The pessimistic view is that "real" reading ability was not altered, since no gains were shown on the more reliable (and, to some, more valid) of our measures. The ability to perform well on the MAT or other commonly used reading achievement tests portends great expectations for a child, at least in future school performance. When MAT is not changed, one could argue, the changes shown on other instruments are either short-term or not really changes in the reading skills tapped by MAT.

We recognize the possibility that our procedure may not produce long-term change. Certainly, we have not proven that it does. However, we cannot accept a claim that we have not affected reading performance. What could possibly be a better indicator of reading ability than measurements of the child's ability to reconstruct afterwards what he has just read? More work is needed before we can be sure how widespread and permanent is our effect. The effect itself is difficult to deny. It may, however, involve only a small component of reading skill.

We briefly mention several unresolved questions that are left for the replication of this study (which is now underway). First, if the MAT posttest is given with imagery instructions, will there be a training effect on it? Second, is it certain that the skills of those in the experimental group improved or did the skills of those in the control group deteriorate? In the replication, use of particular passages for paraphrase recall will be counterbalanced over the pretest and two posttests so that we can look at a temporal trend as well as differences between conditions. Finally, reading times will be taken in the paraphrase recall test so that we can ascertain whether better performance comes merely from longer study in the self-paced test.

E. Mental Imagery and Comprehension in First Graders: Experiment 5

The arguments we developed in Section I. C. suggest that preoperational children should not benefit from the imagery procedure in prose learning if they have an inability to maintain and coordinate simultaneous representations of the overall passage content, the spatial arrangement of text on the page (or the temporal arrangement of spoken words), and the information necessary to parse the sentence currently being processed.

For a variety of technical reasons, we have not been able to run a reading comprehension study on children so young. However, the same basic arguments should apply to listening comprehension, though not so strongly. That is, in listening to a continuing message, presented at a rapid pace, one must be both linguistically processing what is being heard at any instant and maintaining a foreground (context) of what has been heard up until then. Further, a visual organization oriented around a cartoon or constructed image may not be easily coordinated with the verbal representation originally generated. Finally, the preoperational child tends to allow previous thoughts to come to mind in an uncontrolled manner. These thoughts, rather than the passage content, may be recorded in the child's image and remembered instead of (or, in interference with) the passage information.

1. Method

Materials. Five single-episode stories of 30-75 words were prepared at a level easily understandable to first graders. Each was recorded on tape. For each story, a background scene and a set of cut-out objects were prepared such that every action of the story could be illustrated with some of the cut-outs on the background. Two of the stories shared the same background. The backgrounds and cut-outs were in color and were laminated.

Subjects. Forty-seven first-grade students participated in this experiment. Twenty-three of them came from a campus laboratory school, while the remaining 24 came from an urban public school composed almost entirely of children from lower middle class black families. Subjects from the campus laboratory school were randomly assigned to the picture or the control conditions. For reasons relating to operating procedures at the urban public school, picture subjects came from one of two parallel classes while control subjects came from the other. (Analysis of these students' standardized achievement test scores showed no differences between the classes.)

Procedure. Subjects were run individually. The basic experimental scheme is shown in Figure 1 (p. 22). In the picture condition, the subject was asked to listen to the five stories in turn. After hearing a story, the subject would select a background and some cut-outs, and illustrate the story content. Subjects were not given imagery instructions explicitly, but they were told in advance that they would be required to illustrate what happened in each story by making a picture with the cut-outs. The illustrations were photographed for later analysis, and then the pieces were put back into a standard arrangement with the others. All of the backgrounds and cut-outs were visible to the picture subjects throughout the experiment. After a control subject finished hearing a passage, he spent an equivalent amount of time (two minutes) doing simple geometry problems (e.g., given a picture of a square and a picture of a circle, draw a picture of the one superimposed upon the other). After hearing every story and doing the picture or control task after each, the subject was given a clue for each story in turn and asked to retell the story without regard to its exact wording. The subject's recall was recorded on tape. Participants in both groups were told in advance that they would be asked to recall the stories later and, therefore, that they should try to remember them carefully.

Scoring. Scoring of predication recall was as described in Section II.A.1, except that responses were classified as either verbatim, correct meaning but not verbatim (synonym), incomplete, overly specific, or not recalled. For example, if (1) were in the original passage, then the same wording in a recall protocol would be scored as verbatim recall of the underlying predication of (1). Sentence 2 would be scored as synonymous recall, (3) would be incomplete specification, and (4) would be overspecification.

(1) The officer patched the roof.

(2) The officer repaired the roof.

(3) The officer fixed the house.

(4) The officer replaced some shingles.

In this experiment, we scored protocols two other ways: noun recall and main idea recall. For noun recall scoring, the same basic categories were used, except that incomplete recall meant recall of a superordinate category (e.g., soldier instead of officer), while over-specified recall meant recalling an instance of the correct category (e.g., lieutenant instead of officer).

Recall of the main idea was crudely scored on a five-point scale, in which the scorer subjectively rated main idea recall adequacy.

An attempt was also made to score the constructed pictures of experimental subjects. A five-point scale was used in which one point was given for each of the following criteria: (a) use of the correct background scene; (b) appropriate placement of items on the background (e.g., monkeys on monkey island rather than on the picnic table); (c) at least one character or object from the story illustrated with a cut-out figure; (d) no irrelevant cut-outs used; and (e) reasonable representation of the main idea of the story.

2. Results

Table 8 shows predicate and noun recall in each of the four categories plus the mean total numbers of predicates and nouns recalled even partially. The equivalent of the measure used in the other experiments is correct meaning recall. This is the sum of verbatim and synonymous recall scores. Summary means for the verbatim-plus-synonymous scores are shown in Table 9.

TABLE 8

Mean Percentage of Recall by Adequacy Categories

School	Condition	Verbatim	Synonymous	Incomplete	Over Specified	Total
Predications:						
Lab	Picture	18.6	12.1	5.8	1.7	38.4
Public	Picture	10.4	9.8	4.8	0.9	25.9
Lab	Control	24.6	14.6	6.4	2.2	47.8
Public	Control	18.9	10.1	5.3	1.1	35.4
Nouns:						
School	Condition	Verbatim	Synonymous	Incomplete	Over Specified	Total
Lab	Picture	45.8	3.6	3.2	0.8	53.4
Public	Picture	31.7	4.8	5.2	0.2	41.9
Lab	Control	55.4	4.2	5.3	1.4	66.3
Public	Control	49.0	4.2	5.3	0.5	59.0

TABLE 9

Summary of Percentage Means

Verbatim or synonymous predication recall:

	Picture	Control	Combined
Lab School	30.7	39.2	34.4
Public School	20.2	29.0	24.6
Combined	25.6	33.6	

Verbatim or synonymous noun recall:

	Picture	Control	Combined
Lab School	49.4	59.6	53.8
Public School	36.6	53.3	44.9
Combined	43.2	56.1	

Main idea adequate:

	Picture	Control	Combined
Lab School	60	86	72
Public School	38	66	52
Combined	50	76	

Analyses of variance were performed on these scores, using School and Condition as between-subject independent variables. All School by Condition interactions were not significant, $F < 1$. At all levels of scoring rigor, the campus lab school did better than the urban public school, and the picture condition did worse than the control. The specific statistics are reported in Table 10.

TABLE 10

F-Statistics for Analyses of Variance
(All with 1 and 43 degrees of freedom)

Category	Level of Criterion	F-Statistics Conditions	Schools
Noun recall	Verbatim	6.55*	11.3**
	Correct meaning	5.16*	10.2**
	At least partial meaning	4.43*	11.6**
Predicate recall	Verbatim	8.65**	9.42**
	Correct meaning	8.62**	6.03*
	At least partial meaning	10.4 **	6.90*

*p < .05
**p < .01

Another item of some interest is the crude rating of whether or not the main idea of the passage was recalled. It is possible for subjects in different treatments to all get the main idea but to differ in whether or not they recall details. That did not happen in the present study. Specifically, for each subject a score of zero to five was assigned, dependent upon the number of stories from which the main idea was judged to be adequately recalled. An analysis of variance of these scores (shown as percentages in Table 9) revealed a schools effect, $F(1, 43) = 5.61$, $p < .05$, a conditions effect, $F(1, 43) = 9.86$, $p < .01$, and no interaction, $F < 1$.

The final set of results concerns the properties of the actual pictures that subjects in the picture condition constructed. Each subject's picture for each story was photographed and later rated as described under Scoring above. The campus lab school tended to have more pictures of more subjects meeting the various criteria than the urban public school did, as shown in Table 11. The differences only approach significance, except for the large difference (85 percent versus 48 percent) on whether or not the main idea of the story was represented in the picture, $\underline{z} = 1.96$, $\underline{p} \leq .05$. The scores can be cumulated by giving each subject one point for each of the five criteria passed on each of the five passages. The resulting picture scores are significantly higher for the campus lab school than for the urban public school (means = 23.2 versus 18.8, $\underline{p} < .05$), but this is really due mainly to the single criterion of whether the main idea is captured or not. The picture quality scores correlate 0.51 with the partial-or-complete predication recall measure ($\underline{p} < .01$).

TABLE 11

Data on Adequacy of Subjects' Pictures

	Correct Background	Items Placed Correctly	Some Relevant Items	Number of Irrelevant Items	Main Idea
Campus School (N = 13):					
Mean Percent Adequate	100	89	100	89	85
Number of Subjects Adequate on All Five Stories	13	7	13	8	4
Urban School (N = 12):					
Mean Percent Adequate	92	72	97	63	48
Number of Subjects Adequate on All Five Stories	9	4	11	2	1

3. Discussion

The present results agree with the hypothesis stated above; the preoperational children are hindered by the imagery strategy rather than helped. We suggest that this is because they cannot coordinate the organization of their temporary representations of temporal word order, small-text unit parsing and interpretation, and passage semantic structure. Thus, the added processing demand actually interferes with even the low level of comprehension of which these younger children are capable.

There is a test of this hypothesis, which we are currently carrying out in collaboration with Joel Levin at the University of Wisconsin-Madison. One can provide the surface-structure organization for the child by presenting one sentence at a time for listening followed by illustration. This procedure will eliminate the possibility that thinking about the "cartoon" interferes with keeping track of the order of new words as they come in. If each new glance at the background picture redintegrates what was seen there before, then this sentence-by-sentence presentation may even show a positive effect for the illustration condition, since more organized representation of the content would result. We await the empirical verification of this speculation.

Further converging evidence for the basic hypothesis underlying this experiment comes from the picture adequacy scores. Depending upon the particular criterion (see Table 11), picture adequacy is between 85 percent and 100 percent for the campus lab school subjects in the picture condition. However, their control counterparts still do about 30 percent better on the various recall measures. This suggests that picture adequacy is not sufficient for memory adequacy. However, within the picture group, picture adequacy predicts recall adequacy.

These two results are not inconsistent if viewed in the context of the preoperational mediation-interference hypothesis we have advanced and with the results of Section II. C. in mind. From the data on the older children, we see that picture instructions without extended imagery training do not produce improved memory in these children with recently acquired (on the average) operational thinking ability. Thus the correlation between illustrating and remembering cannot reflect picture facilitation of memory.

We suggest that picture adequacy is a measure of the extent to which the extra picturing task is a source of interference with memory operations. When the extra picture task strongly interferes with the memorizing task, then both performances are deteriorated. On the other hand, for some subjects, there is less of an interference problem; hence, scores on both pictures and recalls are higher. The explanation is partially ad hoc, but the nonfacilitative role of pictures even in older children from the same school is shown separately in Section II. C.

F. General Summary of Results

1. Developmental Trend

While the experiments are not perfectly compatible, they strongly suggest a developmental trend in ability to profit from the imagery/illustration task. Six-year-olds are inhibited in their performance by the extra task demands, nine-year-olds are able to benefit from the illustration only after rather specific training, and adults benefit spontaneously. This is a very different trend from that reported for prompted use of imagery in paired associate learning (cf. Rohwer, in press). When instructed to do so, even six-year-olds use imagery (or picturing) in learning paired associates, and they benefit from doing so.

The present results show a trend more like that found by Rohwer for spontaneous (i.e., uninstructed) use of imagery strategies. This ability seems to develop (in children who are not economically deprived) between the ages of eight and fifteen (cf. Rohwer, in press). Learning of prose and other meaningful sets of information is much more natural for the child than paired associate learning. Thus a child's spontaneous performance in the paired associate task should be influenced by his experience in meaningful learning situations. Our results suggest that in the more complex learning situations children actually face, imagery mediation does not help until about age nine. Not surprisingly, children do not use the strategy spontaneously until they reach that age.

The next step, clearly, is to further break down the comprehension and learning process into specific skills. Once this is done, we can study the development of each of those skills and pinpoint the source of young children's mediation deficiencies. As Pylyshyn (1973) has pointed out, there will be little value in keeping the imagery construct as we build a model of comprehension and learning of prose. However, the various imagery studies carried out thus far present an important range of results to be accounted for by forthcoming process models. As we have stated in Section I, we believe that attentional and working-memory processes are in large part responsible for prose imagery effects. Our future experimental plans are sketched in Section III below.

2. <u>Time-on-Task Controls</u>

Most of our present results can be accounted for by a more superficial hypothesis--that time available for reading and/or thinking about the passages is always greater for the imagery group.[2] In Experiment 1, where time is controlled, the groups do not differ. In Experiments 2 and 3, the imagery group is more able to continue rehearsing the story during the cartoon task than the control group is. This is also true in Experiment 5. Even in Experiment 4, the posttest is self-paced. Our training may have taught children only to read more slowly and carefully. The younger children may suffer in the imagery task only because they have not learned to rehearse that which they wish to remember.

This general argument has some merit. The processes we claim are involved in using imagery to advantage are rehearsal processes. That is, they involve internally-directed changings in working memory. However, they may be specific to particular pieces of the text content. If, as we suspect, the result of imagery mediation is a more connected and organized associative network for the prose content (cf. J. Anderson & Bower, 1973), then it is not just the quantity of rehearsal, but also the quality, that is important.

There are several lines of evidence that answer more directly these arguments about controlling time on task. First, we know from Experiment 3 herein and from several adult studies preliminary to Mary Beth Curtis's master's thesis that (a) imagery groups read prose more slowly, but (b) reading times do not correlate with recall performance. This suggests that "imagery" processing takes time, but that the time would not otherwise be put to good use. In another of Curtis's master's related studies, she replicated Experiment 2 with a fixed reading period rather than a subject-controlled one. During that

---

[2] We thank James Voss for pointing this out.

period, subjects were to read through the passage as many times as they wished. They flicked a micro switch each time they came to the end of the passage. Curtis found that subjects reading under the cartoon instructions read the passage more slowly the <u>first</u> time through. The basic results of Experiment 2 replicated, even though <u>total</u> reading time was constant.

Another time-on-task problem is raised by the problem of the cartoon drawing itself. Even after the subjects finish reading, those in the cartoon group can rehearse the story while drawing the cartoon. The control groups have other things to think about. We plan to replicate some of our results with the cartoon coming after the recall task as a means of assuring that our effects are primarily due to processes that occur while reading. Meanwhile, another study we have yet to report provides some evidence.

As we will discuss in Section III, we have conducted a few experiments in which cartoon tasks and imagery instructions have been replaced by instructions to keep in mind the main points of the passage while reading and to write a verbal summary of the main points rather than draw a cartoon. The results are identical to those with cartoons. In one of the verbal summary experiments, we replaced the summary task with a passage alphabetizing task half of the time. Thus, these subjects were told to always be ready for the summary task, but they only did the summary half the time. A second group was given summary-emphasizing instructions and summarized every time, and a third group knew in advance that they would be required to alphabetize every time. In this experiment, there was also a fixed three-minute reading period for each passage. Thus, the experiment is a variant of our Experiment 2 that controls reading time and content-related time. However, this experiment uses a task that we only hypothesize to be equivalent to the cartoon task. The results tend to refute simple study-time explanations of our other experiments. The group that received main-idea instructions and always wrote summaries recalled 19 percent of the predications in their passages. The all-alphabetizing group recalled only 10 percent, significantly lower than the other two conditions ($p < .05$). The half-and-half group, which read as if a summary would always be required recalled 22 percent when the summary was required and 19 percent (not significantly different) after an alphabetizing task.

Thus, time-on-task interpretations of our results appear to be wrong. However, this problem must still be kept in mind in designing future experiments.

III. Implications for Future Research

A. Treatment-by-Structure Interaction Studies

A general theme of this paper has been that process models must now be proposed to account for the cognitive activities of subjects whose recall changes as a result of imagery instruction. Then, experiments will need to be carried out to test those models. We can at least sketch one kind of experiment that we have been trying to carry out.

The rough model we propose is the following. In addition to an immediate sentence decoder, the subject maintains a buffer (Atkinson & Shiffrin, 1968) or working memory in which certain propositions temporarily reside after being decoded. The subject necessarily divides his mental effort between decoding and buffer manipulation. Recall depends (a) upon the propositions being decoded, and (b) upon the formation of interpropositional connections which provide the "organization" necessary to recall. We further assume that retrieval of the prose content from memory will be more complete if the semantic structure of the prose is congruent with the memory structure of inter-propositional associations. We believe that a variation on J. Anderson and Bower's HAM model (Human Associative Memory, 1973) can be produced which realizes these assumptions.

We wish to model the effect of imagery instructions as an increased allocation of effort to the working-memory maintenance functions. Intuitively, it seems reasonable to hypothesize that keeping in mind a picture of what is happening in a story can be represented as greater allocation of working memory to key propositions of the passage rather than to the most recently decoded propositions. This kind of selective filling of the working-memory buffer may be one "higher level of processing" that facilitates long-term memory.

One way to test this idea is to look at the effects of imagery instructions on learning of prose with varying correlations between semantic structure and sentence (proposition) order. Some passages have propositions so ordered that propositions which need to be associated in memory are contiguous in the prose. In this case, selective allocation of propositions to working memory should not produce better results than a strictly ordered, last-in-first-out strategy. Thus, imagery instructions should fail to facilitate learning of such prose. On the other hand, less sequential prose should benefit from the imagery instructions, since association of each proposition with its neighbors will not be sufficient to permit easy retrieval.

Curtis has been trying to produce an interaction between imagery instructions and passage structure with only moderate success. We now think that her lack of success resulted from inadequate difference between passage structures--both were fairly amenable to nonimagery processing. We will try this task again with better control of the passages. Meanwhile, we have learned of a experiment by Sandra Koser (personal communication) at Cornell which is related to our work. She compared name-organized passages with attribute-organized passages (cf. Frase, 1972) and found that imagery instructions facilitated performance only on the attribute-organized form. That form is usually more poorly learned.

Inspired by Koser's success, we plan to continue instruction-by-structure interaction studies until we can control the imagery effect.

B. Imagery versus Verbal Mediation

The previous arguments seem very independent of any need for a special visual-spatial processor that might be activated by imagery instructions. We spoke only of working memory allocation, not of visual process. Consequently, it should be possible to produce an identical effect with verbal processing instructions. In particular, keeping in mind a picture of what is going on in the passage so that one could draw a cartoon should be equivalent to keeping in mind the main ideas of a passage so that one could write a verbal summary. Preliminary work in our laboratory indicates that this is the case, at least for a gross amount of paraphrase recall. We expect to examine the particular propositions recalled in the two conditions in an attempt to determine whether imagery instructions produce more recall of picturable information than verbal instructions.

C. Structural Analysis of Recall Protocols

Going beyond the research proposed in III. A. above will require an operational definition of the key propositions that we think are kept in working memory under imagery instructions. One possibility is the kind of logical analysis suggested by Crothers (1972), and others. However, such analyses have two problems: (a) they are hard to do; and (b) they substitute a priori hypotheses about prose structure in memory for empirically-verified knowledge, leading to a rather excessive number of not-well-established theoretical assumptions. Experimental tests that fail then have dubious information value, because it is not clear whether the semantic structure theory or the working memory allocation theory is deficient. To avoid this problem, we are attempting to specify empirically the optimal memory organization for prose.

The work is only now beginning, but we can briefly sketch our plans here. In a number of recent experiments involving paraphrase recall of prose, we have kept detailed records of which propositions are actually recalled by which subjects. The data are stored as a matrix of ones and zeroes where the i, j cell is one if subject i recalled proposition j. Let us call this matrix of data D. The co-occurrence matrix

$$C = D^T D$$

can then be computed. The cell c_{ij} contains the number of subjects who recalled both proposition i and proposition j. Propositions that co-occur frequently can be thought of as highly interconnected in the memory scheme common to most subjects. Unfortunately, two propositions may co-occur less often even though they are highly interassociated if recall of both of them is contingent upon recall of a more basic key proposition. So, we build a similarity matrix S, that scales each c_{ij} by the maximum it could have assumed:

$$S_{ij} = \frac{c_{ij}}{c_{ii} + c_{jj} - c_{ij}}$$

The S matrix is then subjected to various clustering analyses. We can then define a well-organized passage as one in which clusters of propositions, as derived from S, are presented together. Such passages should not be helped much by imagery. On the other hand, passages that have interproposition distances (perhaps defined as number of intervening clauses) that do not correlate with corresponding S values should be more facilitated by the imagery-type instructions. We will be testing this hypothesis shortly.

References

Anderson, J. R., & Bower, G. H. Human associative memory. Washton, D. C.: V. H. Winston, 1973.

Anderson, R. C. Encoding processes in the storage and retrieval of sentences. Journal of Experimental Psychology, 1971, 91, 338-341.

Anderson, R. C., & Hidde, J. L. Imagery and sentence learning. Journal of Educational Psychology, 1971, 62, 526-530.

Anderson, R. C., & Kulhavy, R. W. Imagery and prose learning. Journal of Educational Psychology, 1972, 63, 242-243.

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In J. T. Spence & G. H. Bower (Eds.), Psychology of learning and motivation. Vol. 2. New York: Academic Press, 1968.

Baddeley, A. D., & Warrington, E. K. Memory coding and amnesia. Neuropsychologia, 1973, 11, 159-165.

Bartlett, F. C. Remembering. London: Cambridge University Press, 1932.

Bobrow, S. A., & Bower, G. H. Comprehension and recall of sentences. Journal of Experimental Psychology, 1969, 80, 455-461.

Bower, G. H. Mental imagery and associative learning. In L. W. Gregg (Ed.), Cognition in learning and memory. New York: Wiley, 1972.

Bransford, J. D., Barclay, J. R., & Franks, J. J. Sentence memory: A constructive versus interpretive approach. Cognitive Psychology, 1972, 3, 193-209.

Chafe, W. L. Discourse structure and human knowledge. In J. B. Carroll & R. O. Freedle (Eds.), Language comprehension and the acquisition of knowledge. Washington, D. C.: V. H. Winston, 1972.

Clark, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 335-359.

Crothers, E. J. Memory structure and the recall of discourse. In J. B. Carroll & R. O. Freedle (Eds.), Language comprehension and the acquisition of knowledge. Washington, D. C.: V. H. Winston, 1972.

Dooling, D. J., & Lachman, R. Effects of comprehension on retention of prose. Journal of Experimental Psychology, 1971, 88, 216-222.

Dooling, D. J., & Mullet, R. L. Locus of thematic effects in retention of prose. Journal of Experimental Psychology, 1973, 97, 404-406.

Frase, L. T. Structural analysis of the knowledge that results from thinking about text. Journal of Educational Psychology, 1969, 60 (Monograph Supplement 6).

Frase, L. T. Maintenance and control in the acquisition of knowledge from written materials. In J. B. Carroll & R. O. Freedle (Eds.), Language comprehension and the acquisition of knowledge. Washington, D. C.: V. H. Winston, 1972.

Ginsburg, H., & Opper, S. Piaget's theory of intellectual development. Englewood Cliffs, N. J.: Prentice-Hall, 1969.

Levin, J. R. What have we learned about maximizing what children learn? Madison: University of Wisconsin, Research and Development Center for Cognitive Learning, 1974.

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

Matz, R., & Rohwer, W. D., Jr. Visual elaboration of text. Paper presented at the annual meeting of the American Educational Research Association, New York, March 1971.

Olton, R. M. The effect of a mnemonic upon retention of paired-associate material. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 43-48.

Paivio, A. U. Imagery and verbal processes. New York: Holt, Rinehart and Winston, 1971.

Paris, S. G. Children's constructive memory. Paper presented at the meeting of the Society for Research in Child Development, Philadelphia, 1973.

Perfetti, C. A., & Goldman, S. R. Thematization and sentence retrieval. Journal of Verbal Learning and Verbal Behavior, 1974, 13, in press.

Piaget, J., & Inhelder, B. Mental imagery in the child. New York: Basic Books, 1971.

Piaget, J., & Inhelder, B. Memory and intelligence. London: Routledge and Kegan Paul, 1973.

Pylyshyn, Z. W. What the mind's eye tells the mind's brain: A critique of mental imagery. Psychological Bulletin, 1973, 80, 1-24.

Rohwer, W. D., Jr. Elaboration and learning in childhood and adolescence. In H. W. Reese (Ed.), Advances in child development and behavior. New York: Academic Press, in press.

Samuels, S. J. Effects of pictures on learning to read, comprehension, and attitudes. Review of Educational Research, 1970, 40, 397-407.

Simon, H. A. What is visual imagery? An information processing interpretation. In L. W. Gregg (Ed.), *Cognition in learning and memory*. New York: Wiley, 1972.

Varley, W. H., Levin, J. R., Severson, R. A., & Wolff, P. Training imagery production in young children through motor involvement. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, April 1973.