DOCUMENT RESUME

ED 268 514 CS 008 420

AUTHOR Caron, Thomas A.
TITLE Developmental Differences in Sentence Constructivity among School Age Children.
PUB DATE Jul 84
NOTE 89p.; M. Ed. Thesis, Western Washington University. Best copy available.
PUB TYPE Reports - Research/Technical (143) -- Dissertations/Theses - Master Theses (042)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Age Differences; Basal Reading; Child Development; Comparative Analysis; *Developmental Stages; Elementary Education; Grade 2; Grade 4; Grade 6; *Inferences; Learning Processes; Linguistic Theory; *Reading Comprehension; *Reading Processes; *Reading Research; Semantics; *Sentence Structure
IDENTIFIERS Blachowicz (C Z)

ABSTRACT

A study examined the existence in elementary school children of (1) sentence constructivity, (2) developmental differences in constructivity, (3) differences in constructivity across performance levels, and (4) differences after a one- or two-day delay. The study was intended as a partial replication of work by C. Z. Blachowicz (1977-78), which demonstrated the phenomenon of sentence constructivity as part of reading comprehension among children. Subjects were 44 second grade, 54 fourth grade, and 50 sixth grade students from intact classes in a public school in Victoria, British Columbia. The students read a basal story and completed a sentence recognition task. The task required the students to indicate sentences that were literally identical to those in the story, but not those that simply meant the same thing. One or two days later, the recognition task was repeated without access to the story. Results were analyzed for sentence types, grade levels, performance levels, and trials (immediate or delayed). Findings indicated support for constructivity in sentence type analysis, a decrease in errors as grade levels increased, no difference across performance levels, and an increase in constructivity after a delay. The results confirm and extend Blachowicz's findings. A five-page bibliography is provided; appendixes include the story used, the recognition test and the raw data presented in tabular form. (FL)

DEVELOPMENTAL DIFFERENCES IN SENTENCE CONSTRUCTIVITY AMONG SCHOOL AGE CHILDREN

A Thesis

Presented to

The Faculty of

Western Washington University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Education

---

by

Thomas A. Caron

July, 1984

MASTER'S THESIS

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Western Washington University, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of this thesis is allowable only for scholarly purposes. It is understood, however, that any copying or publication of this thesis for commercial purposes, or for financial gain, shall not be allowed without my written permission.

Signature Thomas Hamm

Date Sept. 5, 1984

DEVELOPMENTAL DIFFERENCES IN SENTENCE
CONSTRUCTIVITY AMONG SCHOOL AGE CHILDREN

by

Thomas A. Caron

Accepted in Partial Completion
of the Requirements for the Degree
Master of Education

Dean of Graduate School

Advisory Committee

Chairman

ABSTRACT

This study was undertaken to determine the existence of: I) sentence constructivity among school age children, 2) developmental differences in sentence constructivity, 3) differences in sentence constructivity across performance levels, 4) differences in sentence constructivity after a one or two day delay.

Subjects for the study were drawn from grades 2, 4, and grade 6 in one public elementary school in Victoria, B.C. A basal story was read and a recognition test was completed. After a one or two day delay an identical recognition test was administered to the same subjects. A Newman-Keuls Analysis of Variance was performed across: I) sentence types, 2) grade levels, 3) performance levels, and 4) trials.

Results indicated support for constructivity in sentence type analysis, a decrease in errors as grade levels increased, no difference across performance levels, and an increase in constructivity after a delay.

These results confirm and extend the findings of Blachowicz (1977-78).

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

ACKNOWLEDGEMENTS

It is with a great deal of gratitude that I wish to thank the following people for their generous assistance in this work: Dr. Ted Mork, Dr. George Lamb, Dr. Marvin Klein, Dr. John Towner, all at Western Washington University; Dr. Peter Evans, and Dr. Pat Conkin, at the University of Victoria.

I also wish to thank the Principal and Staff of the Children's school in School District No. 62, in Victoria, B.C., for their unqualified co-operation.

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

Statement of the Problem

The present study is a partial replication of the previous work of Blachowicz (1977-78) in which was demonstrated the phenomenon of sentence constructivity as part of reading comprehension among school age children.

Blachowicz presented students of varying ages with groups of sentences relating spatial ideas followed by a sentence recognition test. This test was comprised of four types of sentences: True and False Premise, and True and False Inference, the latter having been developed from semantically congruent and incongruent associations of related premises, respectively. Although the only test sentences which had actually been presented at acquisition were the True Premise sentences, Blachowicz found that all her students tended to 'recognize' True Inference sentences as if they also had been presented. She suggested this indicated " a pattern of performance which strongly supports the construc-

tive hypothesis " (p.194); that is, students were more concerned with reading and remembering the ideas represented in print than the actual words used to represent those ideas.

Blachowicz cites, as a limitation of her work, the narrow semantic nature resulting from restriction to the inter-relations of a spatial nature in her paragraphs. In addition, it is stated in Blachowicz (1978-79) that the possibility of generalization is inhibited by the limited scope and length of test paragraphs compared to usual primary reading material.

It was the intention of the present study to overcome these limitations of narrow semantic nature and limited scope in the following manner: test paragraphs were drawn from actual primary reading material. Although Blachowicz included adults as subjects, the focus of the present study was on elementary school children. Other factors added in the present study include comparison of good and poor readers and a delayed recognition test.

The sentence recognition test used in the present study consisted of three sentence types. The False Inference category was deemed redundant. It was decided to use fewer sentences after a pilot study using more proved too tiring for the youngest subjects. In keeping with Blachowicz's intent to minimize the difficulty of reading material used, the Inference and False sentences developed used only words which had appeared in the story.

Blachowicz had analyzed patterns of choice by sentence type within and between grades. The purpose of analysis by sentence type within grades was to demonstrate the tendency of all subjects to 'recognize' Inference sentences. Analysis between grades provided a means to determine developmental differences.

For further clarification of the processes that may be active, two additional factors were identified: within each grade error patterns were compared between Good and Poor performing students (as measured by teacher judgement); in addition the sentence recognition test was administered to each

subject a second time, a day or two later, to provide a Delayed score for comparison with Immediate scores of recognition performance immediately following the original reading.

The following null hypotheses were tested:

1. School children (grades 2, 4, and grade 6 in one school in Victoria, B.C.), after reading a "standard basal story" commonly used in grade one, will not demonstrate a tendency to choose Inference sentences in a sentence recognition test.

2. There will be no significant difference between Good and Poor readers (as defined by classroom teacher judgment) in a tendency to choose Inference sentences in a sentence recognition test.

3. There will be no significant difference between grade levels 2, 4, and 6 in a tendency to choose Inference sentences in a sentence recognition test.

4. There will be no significant difference after a one or two day delay in subjects' tendency to

choose Inference sentences in a sentence recognition test.

Expectations and further justifications will be found in the chapter on results.

5.

CHAPTER TWO

Review of the Literature

General Background

The process of reading has long been a subject of serious investigation from Comenius' "Orbis Pictus" in 1657, mentioned by E.B. Huey (1908), to the present day. As Huey implied in his History of Reading, there have always been inter-relationships among culture, mind and symbol. From the earliest pictographs to present day 'word processing', reading has always been characterized by an interaction between the word and the mind that "builds up in the mind an impression of meaningful message." (J.B. Carroll, 1970, p.295).

For a child developing into the adult world of symbols and sense, the context, impact, and meaning of experience changes as the child develops (Piaget, 1952). Action and understanding reach outward from a child's growing perspective through

the continual interplay of development and learning (see Vygotsky, 1962 or Vygotsky, 1978).

Meaning in print surrounds us. The complex activities of writing and reading offer to every child an exciting challenge. Many children meet that challenge, some with difficulty, some without; other children don't seem to find success at all with reading and/or writing, even though virtually every child learns to speak at least one language and many become proficient in several languages. Surely the task of extending our use of language to include reading and writing cannot be a more momentous task than that of learning any language in the first place (Clay, 1979).

What is this challenge of learning to read and write? How can we best teach each child? Such questions have stimulated research and haunted school teachers for years. And many apparent answers often seem to contradict each other (see Chall, 1967). A dichotomy has tended to characterize the fields of research and explanations in the teaching of reading over the years. In practice, research evi-

dence is cited in support of one position or another: for example, Reading as the intelligent use of perception and cognition in the context of natural language development (see Goodman, 1970, or Smith, 1978), or Reading as the end result of sequential skill development (see Gough, 1972, or LaBerge and Samuels, 1974). Evidence has been used to pull together the two opposing views under one new label: such as the Cognitive Clarity Theory (Downing, 1979), or an Interactive Model of Reading (Rumelhart, 1977), or a Progressive Refinement Theory (Collins, Brown, and Larkin, 1980). The list of attempts to define Reading may be as endless as the list of theorists.

The depth and breadth of these powerful and insightful theories are beyond the scope of the present work. Suffice it to say our attempts at clarification necessarily yield more data, more theory, more answers, and more questions. Analysis and synthesis appear increasingly complex and "removed" from the reality of a child learning to read.

More recently have some researchers been sug-

gesting there may not be one 'way' to define reading, and clearly not one way of teaching reading or of learning to read (see Forester, 1976, or Andrews, 1976). There is no 'right' technique or 'correct' method. Learning to read is a complex experience, one of the greatest endeavors and achievements any mind may expect to encounter (Clay, 1979).

One frequent drawback of much reading research is the lack of a definitive relationship between theory and the reality of a most significant population: beginning readers. Relevant theories are not easy to test; it is difficult to imagine an unbiased test of learning to read that uses reading performance as its measure! How much of what we may call 'skilled reading' must actually be acquired through the process of reading? What is it that can be known prior to reading that is not itself reading?

Apart from the bias which must necessarily be introduced by the methodology of a reading experiment, there does seem little doubt that memory for text itself must be a critical factor in an individual's performance in and use of reading; although

specific methodology may influence not only how a task is viewed but also the ways in which subjects remember and subsequently perform.

F.C. Bartlett (1932) portrayed memory for connected discourse as a 're-constructive' mechanism. He believed and attempted to demonstrate that the changes in memory for stories over time are not such as would be predicted by a simple behaviorist 'memory trace' theory. This theory, he suggested, would predict a gradual decay in memory for a story, containing especially errors of deletion and including no more substantive errors in recall than would have existed in the first recollection. Bartlett discovered, on the contrary, that story memory seems to be governed by considerably more personal and interactive factors than the mere objective trace of a recollection. Studying the transcripts of adults' repeated recalls of a story after progressively longer time intervals, he found that subjects tended to alter and restructure the story more significantly over time, adding rather than deleting material, apparently constructing a story in recollection that made more sense to them than had the

original text. "Remembering appears to be far more decisively an affair of construction rather than one of mere reproduction . . . condensation, elaboration, and invention are common features of ordinary remembering, and these all very often involve the mingling of materials belonging originally to different 'schemata'" (p.205). Bartlett defined schema as "an active organization of past reactions and experiences which are always operating in any well-developed organism" (1932, p.201).

Neisser (1967) suggests that not only memory but the act of perception itself is constructive: that the mind, for example, must 'create meaning' rather than simply observe, store, and retrieve it. "Reading is externally guided by thinking. Perhaps we should not be so surprised that it is so poorly understood; we may not understand it until we understand thought itself" (p.136). In this regard Bartlett used the term 'schema' to represent the cognitive structures within which a reader's comprehension is embedded. He suggested it is the influence of these schemata which modifies and reorganizes recall.

II.

Because Bartlett's work was incompatible with much of the psychological research being conducted at the time, it wasn't until comparatively recently that the significance of his work has been brought to light. Considerable research has been conducted with a view to proving or disproving the existence and/or importance of 'reconstructive' memory (see, for example, Barclay, 1973; Gauld and Stephenson, 1967; Golmulicki, 1956; Kintsch, 1973; Zangwill, 1972; among others). Spiro (1975) in a comprehensive analysis of several approaches to story memory cautioned that none of the conditions that would be necessary for 'reconstruction' to occur had actually been present in most memory experiments and this would predetermine the fact that reconstruction was not observed.

In developing a paradigm to test reconstructive theory, Spiro stated, "Reconstructive processes in recall may be demonstrated by taking advantage of known regularities in cognitive structures to systematically predict the kinds of changes in recall that are most likely to occur due to subsequent new inputs to those cognitive structures if recall is in-

deed reconstructive and not tied to fixed life-less traces" (p.19). Spiro designed an experimental test using various delays between reading and recall, two types of story (balanced and imbalanced), and three kinds of ancillary information: consistent, contradictory, and none. Bartlett, he pointed out, had presented a North American Indian legend to British born experimental subjects. The mismatch between the culture from which the story came and that of the experimental subjects would tend to encourage 'reconstruction' in recall if only to 'rationalize' a story which must have appeared strange and incomprehensible.

Spiro's experiment was designed to isolate the 'comprehensibility' of the story read as a factor. Reconstructive theory would predict that the less comprehensible the story the more likelihood for reconstruction in recall. Secondly the provision of consistent, contradictory, or no ancillary information following the reading was designed to control the influences on cognitive structures that might conceivably alter recall. Spiro reported, "all the predictions of Reconstructive Theory were conclu-

clusively and unambiguously confirmed" (p.40).

Selected Related Literature

Bransford and Franks (1971) used a sentence recognition test to study inferences during reading and remembering. After constructing sentences containing four major ideas or propositions each, they presented subjects with sentences made up of one, two, or three of these idea sets. Complete four-idea sentences were never presented during this acquisition phase. Subjects were asked to pay close attention as questions would be asked about the sentences later. Following acquisition subjects were shown a list of sentences and asked to chose which were identical to any they had heard in acquisition. This recognition set consisted of some sentences that had identical syntactic structure but meaning that was false as the result of some minor word change, and some sentences that also had not appeared in acquisition but agreed semantically with the underlying meaning of acquisition sentences. These latter were called 'inference' sentences and included the complete four idea sentences from which acquisition sentences had been derived.

An example of a four-idea sentence Bransford and Franks used is: "The rock which rolled down the mountain crushed the tiny hut at the edge of the woods." The following are examples of single idea sentences derived from this complete sentence: "The rock rolled down the mountain," "The rock crushed the tiny hut," "The hut was at the edge of the woods." Bradsford and Franks predicted that subjects who had heard these single idea sentences and sentences relating two and three ideas would confidently identify the original four idea sentence in the recognition phase as one which had been heard even though it had not been presented in acquisition. It is this 'misrecognition' which the experimenters believed indicate the tendency of subjects to integrate textual information in memory and to demonstrate on retrieval the results of this 'reconstruction.'

Results were consistent with Bransford and Franks' predictions and they cite particularly the evidence of high confidence in choosing longer sentences which had not appeared as indicating that subjects were intent on understanding the interrelationships between sentences rather than simply

processing and storing sentences as they were presented. They discount the possibility that poor memory could have been responsible for the effect as subjects were also confident that they had not heard 'Non-Case' sentences. 'Non-Case' were sentences which had been constructed to use ideas from sentences combined in ways that did not represent the integration of idea sets. For example, given the following two sentences of four ideas: "The tall tree in the front yard shaded the man who was smoking his pipe" and "The girl who lives next door broke the large window on the porch", a 'Non-Case' sentence would be: "The man who lives next door broke the large window on the porch." The mismatch across sentence was intended to represent a form of idea integration which a subject would not demonstrate if idea integration represented truly semantic organization.

Bransford and Franks' interpretation that "subjects, in general, could not discriminate novel sentences consonant with the ideas acquired during acquisition (NEWS) from sentences that they had actually heard during acquisition (OLDS)" (p.348) has

raised some interesting questions. Griggs and Keen (1977) have questioned the test procedure after finding the results could be influenced by the nature of the directions given to subjects. If subjects are told what is to be expected of them and are given a forced choice task in which they are told how many acquisition sentences to choose, experimenters report, contrary to Bransford and Franks' claims, that there is clear evidence of subjects' ability to distinguish Old from New sentences and that this result suggests that particular information is stored in memory. It would seem, however, that the retention of memory for Old sentences does not contradict the possibility and importance of 'inference' during reading and remembering.

Others as well (see Moeser, 1982, for example) have raised serious questions about Bransford and Franks' interpretations and experimental design. It would seem the more specific we can make our directions to subjects the more likely it is that we can pre-determine the results of our experiments. Yet, one must keep in mind the more we contrive an experimental situation the less likely it is that

we can expect to observe the behaviour that a subject would be exhibiting in a more 'natural' or 'real life' reading experience.

Walker and Meyer (1980) reviewed research on knowledge integration. Noting the discrepancies that exist among the various theories of Bransford and Franks (1971), Kintsch (1974), and Hayes-Roth (1979), for example, Walker and Meyer point out there seems to be agreement on the following:

1. the integration of related and contiguous propositions occurs spontaneously during the process of acquiring information, and

2. integrated memory structures formed during acquisition aid in comprehension and facilitate higher order cognitive processing. (p433)

They suggest some of the factors active in such integration include:

- degree of correspondence between related facts.
- physical proximity in text.
- weight in content structure.
- degree of factual material committed to memory.

Accepting the existence of some extent of knowledge integration in the process of reading, these and other factors active in such integration may be difficult to control sufficiently in a sentence recognition experiment. Even more important, particularly regarding the learning and teaching of reading, this raises serious questions about the tendency of children to integrate knowledge from text as they are learning how to read that text. How the particular language skill of reading is encouraged in children must necessarily depend on what it is they are attending to in the process. It is apparent that the variety of attitudes towards and expectations of reading which children bring to the task must influence what they are able to do and how it will be done (see Clay, 1972; Downing, Ayers and Shaefer, 1978; Harste, Burke, and Woodward, 1982). As Barclay and Reid (1974) point out, "For the child efficient classroom learning depends to some extent on his ability to semantically integrate when dealing with spoken and written passages" (p.278).

Paris and Lindauer (1976) studying various age

groups found that, although younger children do also exhibit patterns of inference behaviour similar to older children (Paris and Carter, 1973), this ability does appear to improve with age. Paris and Lindauer suggest that this may be a result of children's growing ability to extend their own thought beyond the information they are given.

Relating this tendency to 'inference' in reading to Piagetian mental operations Prawat and Cancelli (1976) suggest that "conservers are more prone than nonconservers to construct meaning from sentence input because conservers have reached the stage of concrete mental operations" (p.50). Besides cognitive ability, Morris and Bransford (1982) report that "people's abilities to draw inferences from two different premises is strongly affected by their current knowledge base" (p.192). P. David Pearson (1982) suggests that "content factors are the knowledge structures residing in our long term semantic memory that determine how well we understand and integrate a particular text" (p.3). There can be apparently considerable barriers to the design of a reading task that would ensure an identifiable degree of comprehension among a population of beginning readers.

Review of Blachowicz

Following early work in constructivity in memory, Camille Blachowicz (1977-78) explored children's responses to a sentence recognition task. Her subjects were drawn from second, fifth, and seventh grades and a group of adults. Materials consisted of an acquisition set (10 short paragraphs read first) and a recognition set (40 sentences, 4 related to each acquisition paragraph).

Using words representative of second grade reading material, Blachowicz created acquisition paragraphs such as the following:

> The birds sat on the branch.
> A hawk flew over it.
> The birds were robins. (p.192)

All paragraphs conformed to this format of spatial relationships. After being given five minutes to read the acquisition paragraphs followed by a three minute interpolated task consisting of math calculations and physical exercise, subjects were

presented with the 40 recognition sentences arranged in random order. For each paragraph read, subjects were asked to classify each of the following 4 types of sentences as sentences they had (YES) or had not (NO) read.

A true statement (TS) - one identical to a sentence in the acquisition paragraph.

A false statement (FS) - one contradicting a single sentence in the acquisition paragraph.

A true inference (TI) - one that correctly links 2 or more sentences in the paragraph.

A false inference (FI) - one that falsely connects 2 or more sentences in the paragraph. (p.193)

For example, drawing from the sample paragraph already given, the four recognition sentences would be:

The birds sat on the branch. (TS)
A hawk flew under it. (FS)
A hawk flew over the birds. (TI)
A hawk flew under the birds. (FI) (p.193)

Blachowicz was then able to compare subjects' misrecognitions of items never seen before (FS, TI, and FI) with their errors in recognition of sentences that had been in the acquisition set (TS).

Blachowicz reports that an examination of the data gathered on falsely recognized inference sentences (TI) indicates "a pattern of performance which strongly supports the constructive hypothesis" (p.194). The primary result was "the strong tendency for all subjects to 'recognize' the semantically congruent inferences as having been present in the original paragraphs" (p.195). Blachowicz suggests the possibility that 'poor memory' could have produced this effect is discounted by the fact that sentences which contradicted the meaning of original sentences while retaining a similar appearance were not as likely to be 'recognized' as sentences with congruent meanings.

In addition to the limitation which Blachowicz admits is introduced by the use of only spatial relations as the semantic nature of her paragraphs, she cites (Blachowicz, 1978-79) the simple scope and length of paragraphs used compared to usual pri-

mary reading material. Generalizations about inferential behaviour made on the basis of such limitations may be suspect when applied to more conventional reading situations.

Given that even the youngest readers, however, are indeed integrating and inter-relating text as they read and are not simply recording a sequential list of words, there may well be significant implications of Blachowicz' results which apply to beginning readers. She suggests, for example, "perhaps attempts at simplifying prose for young readers work against the natural comprehension process" (Blachowicz, 1977-78, p.197).

The work of Blachowicz raises the question of whether or not a similar effect would be evident among young readers if material were drawn from actual primary reading material that is usually used in the classroom.

In relation to her report that there is a decrease in overall error as grade increases, one can

ask, is this an effect related to grade and age level, or would it be evident as well within the same grade across skill levels? Stanovich (1980) cites contradictory literature on possible process differences between good and poor readers. Is there a difference between good and poor readers' tendencies to exhibit this inference effect in sentence recognition? The answer to such a question could have a serious impact on our perception of the relative functioning of readers indicating differing performance abilities.

Considerable research in delayed recall (see Spiro, 1975, or Walker and Meyer, 1980, for a review) seems to indicate that the combined process of reading and recalling is constructive in nature. This raises the question of whether or not a delay of a day or two between reading and recognition test might have any appreciable influence on the tendency of subjects to recognize 'inference' sentences which had not actually appeared in acquisition.

Summary

Blachowicz did find evidence for semantic

constructivity in reading comprehension among school age children. She found total error in recognition decreased as grade level increased, although the overall pattern of responses remained the same. Blachowicz noted the possible lack of correspondence between sentences used in her study and usual primary reading material.

CHAPTER THREE

Design of the Study

This chapter will include a discussion of the materials used, the subjects chosen, and the procedure followed in the present study.

Introduction

To the extent the present study replicates the work of Blachowicz, a story was read silently by groups of students at varying grade levels. After an intervening math related task, students then completed a sentence recognition test which included 'inference' sentences in the story.

In response to a limitation stated by Blachowicz, a story found in usual primary reading material was used in acquisition in place of groups of contrived sentences. This necessitated the choice of original sentences and the design of inference and false sentences for use in the recognition test. False inference sentences were deemed redundant; this category

was left out.

Since constructivity in comprehension has been most often studied through delayed recall, a time delay between reading and recognition was introduced into the design.

In the light of questions regarding possible process differences between good and poor readers, data was further analyzed across the factor of usual reading performance as defined by classroom teachers.

Materials

A story (Appendix A) was chosen from a Nelson reading text which is slated for optional use in Grade One in Canada. It was decided to use a recognition test of nine sentences, three each in the categories: Actual, Inference, and False (Appendix B).

Original sentences chosen for the category of Actual in the sentence recognition test were:

A1: "I have chicken-pox."

A2: "No," said Grandmother.

A3: "What is that?" asked Jack's father.

It has been suggested that one of the limitations of previous work has been the 'contrived nature' of acquisition sentences which may have led too easily to what may be called inferences, as, for example, in the case of the inference:

" A hawk flew over the birds."

having been developed from a combination of the sentences:

"The birds sat on the branch.", and

"A hawk flew over it.".

Although the relevance of the acquisition sentences may be enhanced through the use of usual school reading material, as Blachowicz suggested, clearly a new difficulty is presented for the experimenter by the necessity of developing reasonable and acceptable inference and false sentences from this text for use in a recognition test. In the construction of inference sentences for this study, reference was made to Warren, Nicholas, and Trabasso (1978). They discuss the construction of inferences from text in a variety of ways including the combination of sen-

tences, substitution of pronouns or synonyms, and implied reference to previous antecedents.

From original sentences following each other closely in the text used, the following inference sentences were created: The manner in which inference sentences on the right hand side were developed from pairs of original sentences on the left hand side will be explained briefly.

Original Sentences : Inference Sentences

"I have chicken-pox."
"You have?" said Ted.
The implication implicit in the second sentence is made explicit in the following inference sentence:

II

"You have chicken-pox?" said Ted.

"Can I get up and play?" asked Jack.
"Can I paint?" asked Jack.
These sentences were combined by replacing 'get up' in the first sentence with 'paint' from the second sentence to create the inference sentence:

I2

"Can I play and paint?" asked Jack.

"Can I call Ted?" asked Jack.

"Yes," said Father.

The meaning implicit in Father's answer is made explicit in the inference sentence:

I3

"You can call Ted," said Father.

In a similar manner original sentences were changed or combined in such a way as to create false sentences for use in the recognition test.

Original Sentences : False Sentences

"Can I call Ted?" asked Jack.

In this sentence 'Jack' was replaced with 'Father' who had not asked such a question to create the false sentence:

FI

"Can I call Ted?" asked Father.

Jack said, "I have little red spots."

"Can I paint?" asked Jack.

3I

In the second sentence 'paint' is replaced with 'have little red spots' so that Jack appears to be asking to have the symptoms of his illness in the false sentence:

F2

"Can I have little red spots?" asked Jack.

"Can I get up and play?" asked Jack.

"You have?" said Ted.

In the second sentence 'have' is replaced with 'can play' so that Jack is made to ask an unlikely question in the false sentence:

F3

"You can play?" said Ted.

All sentences in the recognition test used only words appearing in the acquisition story. An attempt was made to keep the inference and false sentences visually and syntactically representative of original sentences. Although only Actual sentences had appeared in the story, the Inference sentences were designed to be semantically congruent with sentences and ideas in the story, whereas each of the false sentences was designed in such a way as to be semantically incongruent with the story

while retaining syntactic similarity with original sentences.

In the absence of any clear measure of validity for such self-created test sentences, twelve teachers, including those whose classes were used as subjects in this study, were given the following informal task: Each was given individually a copy of the story used and a copy of the recognition test. The intent of the inference and false sentences (given above) was briefly explained to each of these teachers. Each was then requested to identify the category to which each recognition sentence must belong. Every inference and false sentence was correctly identified by every teacher. There was one case in which the Original sentence, "I have chicken-pox," was erroneously identified as Inference instead of Actual. This seemed to be the result of working hastily and did not seem to raise any serious question regarding the quality of Inference or False Sentences.

Subjects

The Principal and Teachers of a public elementary school in Victoria, B.C., agreed to allow their students to participate in this study. Because of availability of subjects the grade levels studied varied somewhat from those studied by Blachowicz (grades 2, 4 and grade 7; grades 2, 4 and grade 6). It was not expected this discrepancy would influence results in any serious way. Prior to the study one classroom of grade two students participated in a pilot study to determine whether any procedural changes might be indicated. Those changes are discussed in the following section on procedure.

Subjects for the study (Ss) consisted of 44 grade two students from three classrooms, 54 grade four students from three classrooms, and 50 grade six students from three classrooms. Ss consisted of the total student population at each grade level studied in the one school. Although no age data was collected, all students were within one year of the usual age for each class, that is, 7 years,

9 years, and 11 years old, respectively.

To establish reading performance levels for good x poor analysis of results, each classroom teacher provided a list of students from each class in rank order of the teacher's estimation of student's reading performance (Bridge and Tierney, 1981).

The three class groups at each grade level studied were of differing sizes. In order to establish rank orderings of reading performance for each grade level data was pooled in such a way that rank orderings from each class group were matched.

Rank orderings were divided into smaller groupings such that each class at any given grade level consisted of an equal number of smaller groups. These 'grouped' rank orderings were then placed in coinciding order.

For example, for three grade two classrooms of 20, 14, and 10 students each, respective rank orderings were combined in the following manner:

Grade 2's:	20,	I4,	IO	
the first	5,	3,	2,	followed by
the next	5,	4,	3,	followed by
the next	5,	4,	3,	followed by
the next	5,	3,	2	

Once all classrooms for each grade level were included in this 'inter-meshed' larger ranking, each grade level group was divided in half to provide a high performance (Good) and low performance (Poor) group for each grade level. These groups called Good and Poor were later used in analysis of results.

Procedure

It was determined from the pilot study that some of the grade two subjects might experience difficulty with the reading and/or recognition task. These tasks were, therefore, administered to smaller groups of 4 to 6 students each at the grade two level; whereas identical tests were administered to full class size groups at the grades four and six levels.

Students were given the story (Appendix A)

to read silently with the direction to read carefully as there were going to be questions about it on completion. To provide a buffer for short term memory students were asked to write on a separate sheet of paper the numbers from 20 backwards to 1. The purpose of this exercise between reading and recognition test was to allow the last sentences read to reach equal status in long term memory so that the recency of their being last read would not influence the likelihood of their being recognized.

It was ensured prior to beginning the reading task that each student understood and could accomplish this mathematical task. The experimenter observed the subjects reading and waited until all the students had completed the counting task. Stories were collected and copies of the sentence recognition task (Appendix B) were distributed. Two versions with different orderings of sentences were distributed such that students sitting beside each other did not have identical copies. This was to allow for the possibility that students might attempt to compare their work with each other.

The experimenter read aloud the directions

appearing at the top of each recognition test sheet and explained further or answered questions to ensure that subjects understood that they were expected to circle numbers of those sentences that were literally identical to sentences in the story and not sentences that simply 'meant the same thing.' On completion of the sentence recognition task, subjects were advised that nothing more was required of them that day but that the experimenter would see them again in a day or two.

A day or two later, depending on the availability of the classes for testing, the sentence recognition procedure was repeated (without access to the acquisition story) to provide data for 'delayed' results.

CHAPTER FOUR

Results

Data tabulated from the pilot study were not included in the study proper. The Newman-Keuls test of analysis of variance (Winer, 1971) was used to test the significance of differences between means in what follows. Each of the four areas of interest identified in Chapter One will be described separately.

For each trial (immediate and delayed) of each subject, the sentences chosen in the sentence recognition test were tabulated in each of the following categories:

Actual: those sentences which had actually occurred in the story.

Inference: those not having appeared in the story but constructed so as to be semantically congruent with sentences or ideas in the story.

False: those sentences not having appeared in

in the story but constructed so as to be semantically incongruent with ideas or facts in the story.

Sentence Type

In general, concurrent with the findings of Blachowicz (1977-78), sentence 'type' arose as a significant factor consistently in both immediate and delayed conditions ($p < .001$). Graphs indicating the pattern of response errors will appear under grade level differences. This result indicates a clear tendency on the part of subjects to demonstrate recognition of Inference sentences as having appeared in the story although in fact they had not.

Grade Level Differences

Although Blachowicz notes a predictable decrease in overall error as grade increases, a comparable result in the present study is obscured by the apparent tendency of grade two subjects to exhibit the lowest error of the three grades in the Actual category, while conforming to the expectations of highest error in the Inference and False cate-

gories. Significance resulted in the Inference ($p < .06$) and False ($p < .02$) categories. Mean errors for each sentence type at each grade level are given in Table I.

TABLE I

MEAN ERRORS IN SENTENCE TYPE CATEGORIES

Grade	N	Sentence Type		
		Actual	Inference	False
2	44	0.68	I.6I	0.43
4	54	0.96	I.37	0.28
6	50	0.94	I.20	0.I6

The pattern of these errors may be more evident from the graph in Figure I.

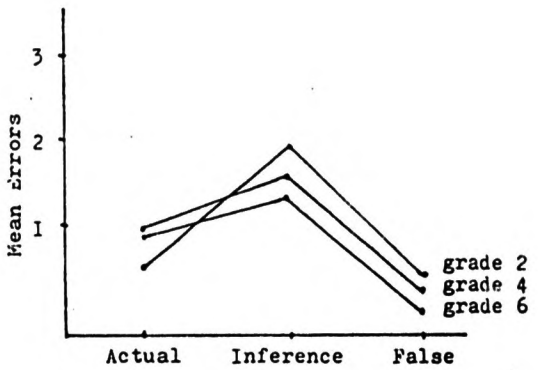

FIG. I. MEAN ERROR BY SENTENCE TYPE

4I .

The patterns of mean error scores (misrecognitions) are similar at each grade studied. Differences among mean errors in Actual sentences did not reach significance. For Inference sentences, mean errors of 1.61, 1.37, and 1.20 for grades 2, 4, and grade 6, respectively were significant at $p<.06$. For False misrecognitions, mean errors of 0.43, 0.28, and 0.16, for grades 2, 4, and grade 6, respectively, also were significant at $p<.02$.

Good x Poor Performance Level Differences

There were no significant differences between good and poor readers (defined by teacher judgment) in mean errors committed in any of the three categories of sentence type: Actual, Inference, or False.

Table 2 specifies the mean errors for each respective category. In the table Performance refers to reading performance level as determined by teachers' rankings divided in half.

TABLE 2

SENTENCE TYPE MEAN ERRORS BY PERFORMANCE LEVELS IN EACH GRADE

Grade	Performance	N	Sentence Type		
			Actual	Inference	False
2	Poor	22	0.64	I.50	0.36
	Good	22	0.73	I.73	0.50
4	Poor	27	I.00	I.56	0.30
	Good	27	0.93	I.I9	0.26
6	Poor	25	I.04	I.20	0.I6
	Good	25	0.84	I.20	0.I6

An analysis of Good x Poor was repeated using the extreme ends of these classifications such that the top and bottom 10 Ss constituted Good and Poor, respectively. No difference between Good and Poor was significant for any sentence type at any grade level.

The means are shown in Table 3.

TABLE 3

SENTENCE TYPE MEAN ERRORS BY REDEFINED PERFORMANCE LEVELS IN EACH GRADE

Grade	Performance	N	Sentence Types		
			Actual	Inference	False
2	Poor	IO	0.50	I.20	0.20
	Good	IO	I.20	I.70	0.40
4	Poor	IO	0.70	I.60	0.20
	Good	IO	I.30	I.20	0.IO
6	Poor	IO	I.IO	I.30	0.0
	Good	IO	0.90	I.20	0.IO

Immediate x Delayed Differences

Differences of mean errors between immediate and delayed recognition tests were significant for all categories of sentence type. With data for all grades pooled, a one or two day delay produced evidence of no significant change in mean errors in Actual, a significant increase of mean errors in Inference ($p < .001$), and a significant decrease of errors in False ($p < .002$) categories of sentence type.

Mean errors for Immediate and Delayed tests for each sentence type are shown in Table 4.

T8 D0 L

TABLE 4

RECOGNITION ERROR MEANS OF IMMEDIATE AND DELAYED TESTS (POOLED GRADES)

Trial	Sentence Type		
	Actual	Inference	False
Immediate	0.87	I.39	0.28
Delayed	0.86	I.72	0.I6

For each grade separately, the results of a one or two day delay between story reading and recognition test were the following:

In grade 2, delay resulted in an increase of mean errors in Actual ($p < .07$), an increase of mean errors in Inference ($p < .00I$), and no significant change of mean errors in False sentence choices.

In grade 4, delay resulted in no significant change of mean error in Actual, an increase of mean error in Inference ($p < .00I$), and a decrease of mean error in False sentence choices ($p < .00I$).

In grade 6, delay resulted in no significant change in Actual mean error, an increase of mean error in Inference ($p < .09$), and no significant change in False mean error.

These results appear in Table 5.

TABLE 5

MEAN ERRORS WITHIN SENTENCE TYPE BY TRIAL

Grade	N	Trial	Sentence Type		
			Actual	Inference	False
2	44	I*	0.68	I.6I	0.43
		2*	0.9I	I.98	0.34
4	54	I	0.96	I.37	0.28
		2	0.8I	I.80	0.09
6	50	I	0.94	I.20	0.I6
		2	0.88	I.40	0.08

* I =Immediate
2 =Delayed

The patterns of error difference between Immediate and Delayed trials are demonstrated more clearly in Figures 2, 3, and 4.

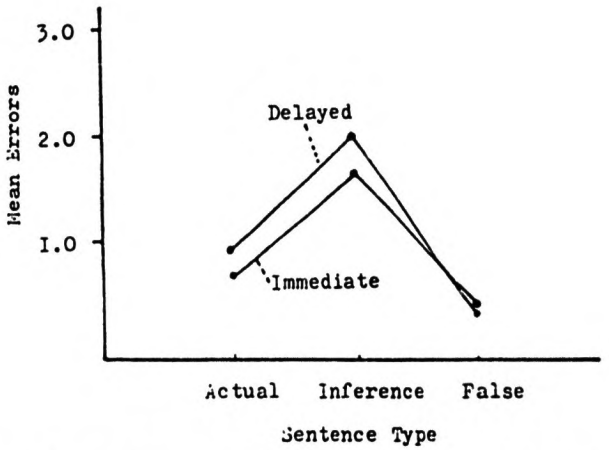

FIG. 2. MEAN ERRORS AT EACH TRIAL BY SENTENCE TYPE (GRADE TWO)

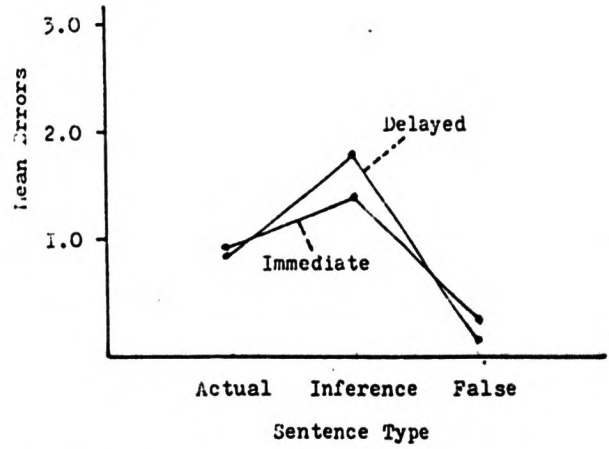

FIG. 3. MEAN ERRORS AT EACH TRIAL BY SENTENCE TYPE (GRADE FOUR)

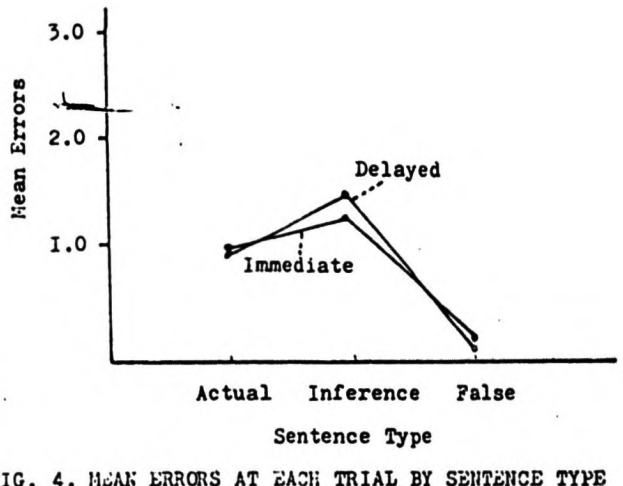

FIG. 4. MEAN ERRORS AT EACH TRIAL BY SENTENCE TYPE (GRADE SIX)

Except for grade two Ss, who exhibited an apparent increase of mean error in Actual, comparison of Immediate recognition test results with results of recognition tests performed a day or two after reading the acquisition story indicate the following for all Ss:

A delay of one or two days between reading and recognition resulted in:

-no significant change in Actual
-an increase in Inference, and
-a decrease in False error means.

CHAPTER FIVE

Discussion

Sentence Type

Given the fact that the directions were clearly stated, "to choose only those sentences which are exactly the same as any appearing in the story", and were followed by the subjects during the sentence recognition tests, we can say there is a difference between the types of errors that can be made:

Actual: An error in this category is represented by the failure to recognize a sentence which did appear in the acquisition story.

Inference: An error in this category is represented by the choice, or recognition, of a semantically congruent sentence which did not appear in the acquisition story.

False: An error in this category is repre-

sented by the choice, or recognition, of a semantically incongruent sentence which did not appear in the story.

The errors in Actual consist of 'non-choices'. This could result from forgetfulness or confusion. It is unlikely that inference errors could be the result of forgetfulness. Bransford and Franks (1971) suggest these errors would indicate 'semantic integration'. It has also been suggested that such errors could be the result of 'semantic confusion' (Hoeser, 1976); however, it is difficult to imagine such confusion not influencing errors among 'semantically incongruent' (False) sentences as well, unless some degree of 'semantic integration' is active. Errors in False sentences might also arise out of difficulty with reading the passage in the first place.

Since sentence 'types' have not been formally standardized, we can use little more than intuition to determine what they actually represent. The sentences used, however, were successfully identified

by category by all twelve teachers at the school where the subjects were obtained. We feel confident, therefore, in saying that the three 'types' do represent distinct categories.

The result that Inference choices were significantly high concurs with the findings of Blachowicz (1977-78) and does suggest that semantic integration is active in subjects' recognition errors.

Grade Level Differences

Each type of recognition error will be considered separately in relation to expected differences.

Actual:

If actual errors are the result of memory loss, we might expect a gradual decline in the incidence of such errors as grade level increases and children may be expected to become more adept at reading and remembering. It is noteworthy that in the present study this does not seem to be the case. All

error scores for Actual were in a similarly low range with one surprise: the lowest Actual error mean was obtained from grade two (Immediate), the group which would be least expected to exhibit the lowest error, although this discrepancy was not significantly different.

Although no difference in Actual error means between grades was significant, it is not unreasonable to speculate that this relatively low score of grade 2's (close to 10 percent probability of F) could have resulted from the altered treatment for this grade level. It will be recalled that a pilot test indicated there might be greater difficulty experienced among the lower grade level students with the tasks in general and that it was decided the grade two level students should be tested in smaller groups (of no more than 6) to obviate their tendencies to either look about themselves with a view to 'copying' neighbours' answers or to randomly choose answers without paying much attention. It was felt that for this group the closer proximity to the experimenter would influence their behaviour in such a way as to ensure greater validity in their results. The exper-

imenter, for example, would be able more easily to circulate among students and provide whatever encouragement and/or assistance might be required to enable students to stay 'on task' and to work independently.

It is certainly possible that the greater attention this group received could have been instrumental in creating a somewhat lower error than expected. In general, however, we must accept that errors in Actual across the three grade levels remained relatively stable and low. There is no evidence in this result of any increase or decrease in 'recognition memory' for sentences as children progress from grade two to grade six.

Inference:

If the choice of inference sentences as having appeared in the acquisition story represents a tendency to base one's choice on inherent semantic ideas rather than literal or syntactic similarity, what can we expect as grade level increases? The work of Blachowicz (1977-78) indicates a fairly consistent

number of errors across all three grade levels: grades 2, 5, and 7 with inference error mean of 4.83, 4.58, and 4.60, respectively, with no clear developmental difference.

The present study, however, did reveal a grade level difference ($p < .06$) for Inference sentence choice (see Table I, in CHAPTER FOUR), the mean errors for Inference being 1.61, 1.37, and 1.20, respectively, for grades 2, 4, and grade 6. Although there have been discrepancies regarding the existence of developmental differences in the use of inferences by children (see Barclay and Reid, 1974; Paris and Carter, 1973, for example) the present results may indicate an increase in precision and care in reading and remembering as grade level increases, which is what one might intuitively expect.

False:

As grade level increases, it would be expected that the tendency to choose False sentences in recognition will decline. Indeed this does seem to be

indicated by the significant decline ($p < .002$) of False error means 0.43, 0.28, and 0.16 through grades 2, 4, and grade 6, respectively. This result could indicate that subjects' clarity about what might have been in the acquisition story increases with grade level. It could also indicate a general improvement of reading behaviour as reading ability increases with increasing grade level. This latter possibility will be considered again in the context of performance level differences.

Good x Poor Performance Level Differences

In view of the results for grade differences one can make certain predictions regarding performance level differences. It seems logical to expect differences in the latter area which are analogous to those found to be significant across grades. Specifically, given there is an indicated decrease in a tendency to choose Inference and to choose False sentences in recognition as reading age increases, might one not expect similar decreases as reading performance increases?

The present study found no such differences (see Table 2). There were no significant differences between good and poor readers (defined by teacher judgment) in mean errors committed in any of the three sentence type categories: Actual, Inference, or False.

On the surface this result is puzzling: although standardized reading tests would indicate clear and often vast differences between performance level in terms of 'reading skills' or abilities, the present recognition test study indicates no demonstrable difference between these groups in either the tendency to err on Actual or False sentences or their tendency to choose Inference sentences in a sentence recognition test. This could imply that readers of varying skill levels exhibit similar tendencies of error and inference in sentence recognition.

Immediate x Delayed Differences

As Spiro (1975) pointed out, a strictly behavioral psychological theory of memory would predict that a delay between reading and recognition would produce no more than a slight increase in error as

a result of memory loss. Within our experimental design, considering only memory loss over a delay, we would predict a slight increase in Actual and False sentence error means and perhaps a decrease in Inference choices as a result of lost 'semantic connections'.

A constructive theory of memory, on the other hand, in which semantic meanings are generalized (Spiro, 1975) would not only predict subjects' tendencies to choose Inference, but would also predict that a delay between reading and recognition would result in an increase in the number of Inference choices and possibly in Actual errors as well. One can conceive of these effects as the result of increasing generalization over time.

In fact the results of the present study do seem to suggest the existence of just such a constructive memory. A one or two day delay between reading and recognition produced in all groups of subjects tested:

- no significant change in Actual error means,
- a significant increase in Inference error means,
- a slight or significant drop in False error means.

These results could suggest that time alone has the effect of "consolidating" what has been read and perhaps enhancing overall understanding.

CHAPTER SIX

Conclusions

Sentence Type

The high incidence of the choice of Inference sentences in recognition among all groups confirms the research of Blachowicz (1977-78) using a sentence recognition test of sentences constructed from a basal story in place of contrived sentences.

Grade Level Differences

The significant decrease in Inference errors as grade level increases may be in contradiction with Blachowicz's results as previously mentioned. One intention in implementing this partial replication was to discover whether there might be any demonstrable differences in children's error patterns resulting from the more familiar story patterns inherent to basals used in schools than found in the contrived logic of spatial relationships previously studied. It was expected that 'semantic integration' would be at least as evident in the basal context

as in the contrived. It appears however that the evidence of semantic integration in recognition may diminish as grade increases or that students at higher grades become more precise. The latter argument would seem to be the more acceptable, although we would also hypothesize that no grade level differences would be evident if the material used were equivalently difficult across grades. The absence of any significant differences between recognition test performances of good and poor readers at any of the grades tested may indicate that what we are calling semantic integration may indeed be an aspect of reading behaviour which is inherent to the process and consistent across all skill and grade levels.

When one considers the consistency of error scores for sentence types across grade and skill levels, one cannot help but consider the possibility that the Inference scores represent some sort of limiting factor: as one grows in reading skill and age one does become more precise in one's recognition memory, but never to the point that the 'misrecognition' of Inference sentences would be completely eliminated. Inference is simply too basic to

thoughtful reading.

In this regard it is interesting to speculate that there may be some kind of balance in skilled reading between 'attention to print' and inference. In the present study, as reading age increased, errors among Actual sentences did not vary, but errors in Inference and False sentences were seen to decrease significantly. Although misrecognition of Actual sentences did not change, the recognition that False sentences were indeed false did seem to improve with reading age, and the tendency to choose Inference sentences did decline as reading age increased.

It appears as if 'skilled' reading might entail the expectation of some loss of specificity, the use of rewording and rephrasing (inference), and clear recognition of 'what does not fit' (False). This conceptualization encourages us to expect some forgetfulness and some inference in skilled reading, as well as the gradual clarification of boundaries (what fits and what does not fit).

It is among the youngest readers that these

expectations may lead us to pay attention differently to reading performance. It may well be that early experience with 'word perfect' expectations of reading performance may actually distract a developing reader from the more flexible interaction between reader and print that more realistically reflects natural reading behaviour.

Performance Level Differences

Superficially, the absence of any clear differences across performance levels on recognition errors is puzzling. One tends to expect better readers to be better at remembering and recognizing as well. Such does not seem to be the case. Again the implication seems to be that there is something less textually specific, less 'word bound' about reading behaviour than we sometimes imagine. It appears there is consistently some small degree of error in Actual recognition, always a somewhat greater degree of Inference error, and a low degree of error in False recognition.

If we think of this as evidence for some kind of 'balance model' of reading, as is suggested, for

example, by Downing's Cognitive Clarity theory (1979), then what kind of behaviour can we justifiably expect of young children learning to read? As reading behaviour develops we must accept that this new experience of identifying ideas in combination of words is going to necessarily involve inference and the re-interpretation of what is drawn from the text.

If it is indeed the case, as implied by results of the present study, that good and poor readers exhibit similar tendencies to error and inference in sentence recognition, this result could have some rather serious implications:

1) Good and poor readers might benefit from a similar approach to teaching reading.

2) Readers (good and poor) even as early as grade two are using inference and would probably benefit from instructional focus on inference in daily reading.

3) The processes, challenges, and benefits of reading comprehension and inference are not limited to 'better' or older readers; to

limit the encouragement of related reading development may well inhibit some readers early in their school careers.

Immediate x Delayed Differences

The results of a one or two day delay between reading and recognition test provide what seems to be the most interesting finding of the present study: rather than what might be logically expected from mere memory loss over time (that is, higher Actual errors, lower Inference errors, and higher False errors), there seems to be evidence for what might be termed a 'consolidation effect': without significant change of errors in Actual, there is a significant increase in Inference error and decrease in False errors.

It may be that time alone, without any verbal intervention relating to the story read, somehow emphasizes the ideas of a story such that readers become more likely to recognize Inferences without any enhanced recognition of Actual sentences and become more able to identify False sentences.

There may be an important implication in this result in support of 'quiet reflection' being at least as important as 'intense study' in enabling a reader to discover deeper meaning. For clarification of the role of reflection, however, one would need to test various combinations and degrees of reflection and study. Our intention in stating this is to emphasize that there is a place for reflection as well as concentrated attention and that this may be a quality and skill which deserves encouragement among children learning to read.

Limitations

This discussion concludes with comments on several limitations and suggested changes that have come to light during the course of this research.

Sentence 'types' could be more definitely identified, more clearly differentiated. Not only would this allow for more precise results but it would also allow researchers to determine more specifically the language processes that may be active. The informal character of the present design leaves open too many questions regarding language processes that may be

influencing results.

A higher number of sentences in the recognition test would be more likely to ensure validity of results. It is possible that, statistically, three sentences per category is insufficient to provide for valid results. Although it is not clear what kind of analogy can actually be made with standardized tests, a good guess by statisticians is that a minimum of 30 items might be required for a test to show clear and valid results.

Regarding the findings related to grade differences, the use of a grade one level story for all three grades tested may have caused differences to disappear as a result of a 'ceiling' effect. It might be advisable to standardize the procedure more equitably by using stories more suitable for each grade level. The use of two stories at each grade level, one being 'easy' the other 'hard' might demonstrate more clearly the existence of any process differences. It would also have aided in relating results to those of other studies if subjects' ages were gathered as data as well as grade levels.

In determining Good x Poor performance comparisons, the use of standardized tests might be an improvement over teacher judgments.

The apparent influence of a one or two day delay is one that is worthy of further research. It would be of benefit to follow delays over a longer period of time and to include an analysis of story recall of identical stories to further clarify recognition results.

Drawing all subjects from only one school may seriously reduce the extent to which one may generalize from results. One must be careful in any case in making generalizations from research in sentence recognition. Several researchers over the years have raised serious questions regarding the validity of sentence recognition design as a methodology. For example, Koeser (1982) and others raise questions regarding the validity of sentence recognition design suggesting that confusion, for example, may be as likely to produce the observed results as may integration.

In general this design does no better than obscure results as controversy over the years has made clear (Spiro, 1975). It is difficult to generalize from recognition results what may be true of recall or how recognition relates to reading. More meaningful data might be gained through studies of free recall or miscue analysis during reading.

It seems evident, however, from the experimenter's experience that some form of a recognition test is enjoyable to children and has the value that even the poorest performing readers seem willing and able to make an attempt at it. This is in contrast to the approach poor performing readers often take to standardized tests. The recognition task is clearly something simple and fun for reader and non-reader alike. It remains only to enhance the validity of recognition results, perhaps through the prudent use of choices. It may be that a more relevant test of recognition items might provide a medium through which both good and poor readers can show their abilities equally well.

BIBLIOGRAPHY

Andrews, N.C. (1976). Six Case Studies in Learning to Read. Unpublished doctoral dissertation, Indiana University, Bloomington. (Xerox Univ. Microfilms No. 77-3274

Barclay, J.R. (1975). The Role of Comprehension in Remembering Sentences. Cognitive Psychology, 4, 229-254.

Barclay, J.R. & Reid, M. (1974). Semantic Integration in Children's Recall of Discourse. Developmental Psychology, 10 (2), 277-281.

Bartlett, F.C. (1932). Remembering: A Study in Experimental and Social Psychology. Cambridge, England: Cambridge University Press.

Blachowicz, C.L. (1977-78). Semantic Constructivity in Children's Comprehension. Reading Research Quarterly, 13(2), 188-201.

Blachowicz, C.L. (1978-79). Factors Affecting Semantic Constructivity in Children's Comprehension. Reading Research Quarterly, 14(2), 165-181.

Bransford, J.R. & Franks, J.J. (1971). The Abstraction of Linguistic Ideas. Cognitive Psychology, 2, 331-350.

Bridge, C.A. & Tierney, R.J. (1981). The Inferential Operations of Children Across Text with Narrative and Expository Tendencies. Journal of Reading Behavior, 13(3), 201-214.

Carroll, J.B. (1970). The Nature of the Reading Process.

In H. Singer & R. Ruddell (Eds.). Theoretical Models and Processes of Reading (pp. 292-303). Newark, DE: International Reading Association.

Chall, J.S. (1967). Learning to Read: The Great Debate. New York: McGraw-Hill.

Clay, M.M. (1979). Reading: the Patterning of Complex Behaviour (2nd Ed.). Auckland: Heinemann Educational Books.

Collins, A., Brown, J.S, & Larkin, K.M. (1980). Inference in Text Understanding. In R.J. Spiro, B.C. Bruce, & W.F. Brewer (Eds.). Theoretical Issues in Reading Comprehension (pp. 385-407). Hillsdale, NJ: Erlbaum.

Downing, J. (1979). Reading and Reasoning. New York: Springer.

Downing, J., Ayers, D., & Shaefer, B. (1978). Conceptual and Perceptual Factors in Learning to Read. Educational Research, 21, 11-17.

Forester, A.D. (1976). The Acquisition of Reading. Unpublished Master's Thesis. University of Victoria, Victoria, B.C.

Gauld, A., & Stephenson, G.M. (1967). Some Experiments Relating to Bartlett's Theory of Remembering. British Journal of Psychology, 58, 39-49.

Golmulicki, B.R., (1956). Recall as an Abstractive Process. Acta Psychologica, 12, 77-94.

Goodman, K.S. (1970). Reading: A Psycholinguistic Guessing Game. In H. Singer & R. Ruddell (Eds.). Theoretical Models and Processes of Reading. (pp. 259-271). Newark, DE: International Reading Association.

Gough, P.B. (1972). One Second of Reading. In J.F. Kavanagh & I.G. Mattingly (Eds.). Language by Ear and by Eye. Cambridge, MA: MIT Press.

Griggs, R.A., & Keen, D.M. (1977). The Role of Test Procedure in Linguistic Integration Studies. Memory and Cognition, 5(6), 685-689.

Harste, J.C.? Burke, C.L., & Woodward, V.A. (1982). Children's Language and World: Initial Encounters with Print. In J.A. Langer & M.T. Burke-Smith (Eds.). Reader Meets Author/Bridging the Gap. (pp. 105-131). Newark, DE: International Reading Association.

Huey, E.B. (1968). The Psychology and Pedagogy of Reading. Cambridge, MA: MIT Press. (Original work published by MacMillan in 1908).

Kintsch, W. (1974). The Representation of Meaning in Memory. New York: Wiley.

LaBerge, D. & Samuels, S.J. (1974). Toward a Theory of Automatic Information Processing in Reading. Cognitive Psychology, 6, 293-323.

Moeser, S.D. (1982). Memory Integration and Memory Interference. Canadian Journal of Psychology, 36(2), 165-188.

Morris, C.D. & Bransford, J.D. (1982). Effective Elaboration and Inferential Reasoning. Memory and Cogni-

tion, 10(2), 188-193.

Neisser, U. (1967). Cognitive Psychology. New York: Appleton.

Paris, S.G. & Carter, A.Y. (1973). Semantic and Constructive Aspects of Sentence Memory in Children. Developmental Psychology, 9, 109-113.

Paris, S.G. & Lindauer, B.K. (1976). The Role of Inference in Children's Comprehension and Memory for Sentences. Cognitive Psychology, 8, 217-227.

Pearson, P.D. (1982). A Context for Instructional Research on Reading Comprehension. (Tech. Rep. No. 250) Urbana-Champaign: University of Illinois, Centre for the Study of Reading.

Piaget, J. (1952). The Origins of Intelligence. New York: International University Press.

Prawat, R.S., & Cancelli, A. (1976). Constructive Memory in Conserving and Nonconserving First Graders. Developmental Psychology, 12(1), 47-50.

Rumelhart, D.E. (1977). Toward an Interactive Model of Reading. In S. Dornic (Ed.). Attention and Performance VI. Hillsdale, NJ: Erlbaum.

Smith, F. (1978). Understanding Reading (2nd Ed.). New York: Holt, Rinehart, and Winston.

Spiro, R.J. (1975). Inferential Reconstruction in Memory for Connected Discourse (Tech. Rep. No. 2) Urbana: Illinois University. Centre for the Study of Reading.

Stanovich, K.F. (1980). Toward an Interactive-compensatory Model of Individual Differences in the Development of Reading Fluency. Reading Research Quarterly, 16(1), 32-71.

Vygotsky, L.S. (1962). Thought and Language. Cambridge, MA: MIT Press and Wiley. (originally published in Russian in 1934).

Vygotsky, L.S. (1978). Mind in Society. Cambridge, MA: Harvard University Press (edited by Michael Cole, Vera John-Steiner, Sylvia Scribner, and Ellen Souberman).

Warren, W.H., Nicholas, D.W., & Trabasso, T. (1978). Event Chains and Inferences in Understanding Narratives. In R.O. Freedle (Ed.). Discourse Processing: Multidisciplinary Approaches. Hillsdale, NJ: Erlbaum.

Winer, B.J. (1971). Statistical Principles in Experimental Design (2nd Ed.). New York: McGraw-Hill.

Text Used For Materials:

McInnes, J.A. (1977). Surprise! Surprise!. Don Mills, Ont.: T. Nelson & Sons (Canada).

APPENDIX A

Acquisition Story Used

"Dad! Dad! Come here!" called Jack.
"What is it?" asked Jack's father.
Jack said, "I have little red spots."
Jack's father looked at the spots.
"You have chicken-pox.
Get into bed," he said.
"Can I get up and play?" asked Jack.
"No," said Grandmother.
"Can I paint?" asked Jack.
"Not in bed!" said Father.
"Can I call Ted?" asked Jack.
"Yes," said Father.
Jack called Ted.
"Hello, Ted," he said.
"I have chicken-pox."
"You have?" said Ted.
"What a surprise!
I have chicken-pox, too.
And chicken-pox is no fun."

APPENDIX B

Recognition Test

DIRECTIONS: Circle the number of each of the following sentences that is exactly the same as any you read in the story.

1. "Can I call Ted?" asked Father.
2. "I have chicken-pox."
3. "No," said Grandmother.
4. "Can I have red spots?" asked Jack.
5. "You have chicken-pox?" said Ted.
6. "What is it?" asked Jack's father.
7. "You can play?" said Ted.
8. "Can I play and paint?" asked Jack.
9. "You can call Ted," said Father.

APPENDIX C

Raw Data

TABULATION OF SUBJECTS' ERRORS IN EACH SENTENCE CATEGORY FOR EACH TRIAL

Grade	Case[a]	Sex	First Recognition Sentence Type			Second Recognition Sentence Type		
			SumA1[b]	SumI1[c]	SumF1[d]	SumA2[e]	SumI2[f]	SumF2[g]
2	001	F	2	3	1	1	3	0
	002	F	0	2	1	2	2	0
	003	F	1	1	1	1	0	0
	004	F	1	1	0	1	2	0
	005	M	1	3	1	0	3	1
	006	F	2	2	0	3	3	1
	007	M	2	2	0	2	2	0
	008	M	1	0	0	1	2	0
	009	F	1	1	0	2	1	0
	010	F	1	2	0	1	3	0
	011	F	0	2	1	0	2	0
	012	F	0	3	1	0	3	0
	013	M	2	3	0	1	3	0
	014	F	0	1	1	1	1	2
	015	F	1	0	1	1	1	0
	016	M	0	2	1	0	2	0
	017	M	0	3	2	0	2	3
	018	F	0	2	0	0	3	0
	019	M	0	1	0	2	2	0
	020	F	1	1	0	1	1	0
	021	M	0	2	0	1	2	0
	022	F	0	1	0	0	1	0
	023	M	1	2	0	1	2	0
	024	F	1	1	0	1	1	0
	025	F	1	3	1	1	3	1
	026	F	1	1	0	1	2	0
	027	F	1	3	1	0	3	2
	028	F	1	3	0	1	2	0
	029	F	1	1	1	1	1	0
	030	F	1	1	1	1	1	0
	031	F	0	2	0	2	2	0
	032	M	0	1	0	1	2	0
	033	M	1	2	1	1	3	2
	034	F	0	1	1	0	3	1

Grade	Case	Sex	First Recognition Sentence Type			Second Recognition Sentence Type		
			SumA1	SumI1	SumF1	SumA2	SumI2	SumF2
2	035	M	0	1	0	0	1	0
	036	F	0	1	0	0	2	0
	037	F	0	2	1	2	2	2
	038	F	0	2	1	0	3	0
	039	F	1	1	0	1	1	0
	040	F	1	1	0	1	2	0
	041	M	0	1	0	0	2	0
	042	F	1	0	0	2	1	0
	043	F	2	1	0	1	1	0
	044	F	0	2	0	1	3	0
4	045	F	2	1	0	1	1	0
	046	F	1	2	0	0	3	0
	047	M	1	2	0	1	2	0
	048	M	1	0	1	1	0	0
	049	F	0	1	0	0	2	0
	050	F	1	0	0	1	1	0
	051	F	0	2	0	0	2	0
	052	F	2	2	0	1	2	0
	053	F	3	2	0	1	2	0
	054	F	2	0	0	2	1	0
	055	F	1	2	0	1	2	0
	056	F	2	2	0	1	2	0
	057	M	0	2	1	0	3	0
	058	F	0	2	0	0	2	0
	059	F	1	0	0	1	0	0
	060	M	0	1	0	0	1	0
	061	M	2	1	1	1	3	0
	062	F	2	1	0	1	2	0
	063	F	1	1	0	3	2	0
	064	M	0	1	1	1	1	0
	065	F	0	2	1	1	3	1
	066	F	0	1	1	0	2	0
	067	F	1	0	0	2	1	0
	068	F	2	0	0	1	2	0
	069	F	0	1	0	0	1	0
	070	M	0	2	0	0	3	0
	071	M	0	1	1	1	3	1
	072	F	1	0	1	1	0	1
	073	F	2	3	1	2	3	1
	074	M	0	1	0	0	1	0
	075	M	1	3	1	0	1	0
	076	M	1	2	0	0	0	0
	077	F	2	2	0	1	2	0
	078	M	2	1	0	2	2	0
	079	F	1	2	0	2	2	0
	080	F	1	2	0	1	3	0
	081	F	0	0	0	0	2	0
	082	F	2	2	1	2	2	0
	083	M	2	1	1	1	1	0

Grade	Case	Sex	First Recognition Sentence Type			Second Recognition Sentence Type		
			SumA1	SumI1	SumF1	SumA2	SumI2	SumF2
4	084	F	1	1	0	1	2	0
	085	M	1	2	1	1	3	0
	086	M	1	0	1	0	3	0
	087	M	2	2	0	2	2	0
	088	M	1	1	0	0	1	0
	089	F	1	2	0	2	2	0
	090	F	2	2	0	1	1	0
	091	M	1	1	0	1	1	0
	092	M	0	1	0	0	3	0
	093	F	1	3	0	1	3	0
	094	F	0	0	0	0	1	0
	095	F	0	3	0	0	3	0
	096	F	1	1	0	1	1	0
	097	M	1	2	1	0	1	1
	098	M	0	1	1	0	2	0
6	099	F	2	2	0	1	1	0
	100	F	1	1	0	2	0	0
	101	F	1	2	0	2	3	0
	102	M	0	0	0	0	0	0
	103	F	1	1	0	1	1	0
	104	M	2	0	0	2	0	0
	105	F	1	2	1	1	2	1
	106	M	1	1	0	1	1	0
	107	F	0	2	0	0	2	0
	108	F	0	1	0	0	1	0
	109	M	1	0	0	3	2	0
	110	F	0	1	0	0	2	0
	111	F	1	1	0	0	1	0
	112	F	1	1	1	1	2	0
	113	M	1	1	0	1	1	0
	114	M	1	1	0	2	1	0
	115	M	0	2	0	1	2	0
	116	F	2	2	1	0	1	0
	117	M	1	1	0	1	2	1
	118	F	1	0	0	0	0	0
	119	F	2	1	0	0	0	0
	120	F	1	3	0	1	2	0
	121	M	0	1	0	0	1	0
	122	F	0	2	1	0	3	1
	123	F	0	1	0	0	1	0
	124	F	1	2	0	1	1	0
	125	M	0	1	0	0	1	0
	126	M	1	0	0	1	0	0
	127	M	0	1	0	1	1	1
	128	F	0	2	1	0	1	0
	129	M	1	2	0	1	2	0

Grade	Case	Sex	First Recognition Sentence Type			Second Recognition Sentence Type		
			SumA1	SumI1	SumF1	SumA2	SumI2	SumF2
6	130	M	2	1	1	2	2	0
	131	M	2	0	0	2	0	0
	132	F	0	2	0	0	2	0
	133	F	2	2	0	2	2	0
	134	F	1	2	1	2	3	0
	135	M	1	0	0	2	0	0
	136	F	0	1	0	0	1	0
	137	F	1	0	1	2	2	0
	138	F	3	1	0	1	2	0
	139	F	2	2	0	2	2	0
	140	F	2	1	0	1	3	0
	141	F	1	1	0	1	2	0
	142	F	1	2	0	1	2	0
	143	M	1	1	0	0	0	0
	144	M	0	0	0	0	2	0
	145	F	1	2	0	0	2	0
	146	M	2	0	0	2	1	0
	147	F	1	2	0	0	2	0
	148	[illegible]	0	2	0	0	2	0

a- Ss within each grade are ranked in descending order of usual reading performance.

b- SumA1 = Sum of errors in Actual (immediate).

c- SumI1 = Sum of errors in Inference (immediate).

d- SumF1 = Sum of errors in False (immediate).

e- SumA2 = Sum of errors in Actual (delayed).

f- SumI2 = Sum of errors in Inference (delayed).

g- SumF2 = Sum of errors in False (delayed).