DOCUMENT RESUME

ED 207 019 CS 006 266

AUTHOR	Wilkinson, Alex Cherry
TITLE	Children's Understanding of Written and Spoken Discourse. Final Report.
INSTITUTION	Wisconsin Univ., Madison. Dept. of Psychology.
SPONS AGENCY	National Inst. of Education (ED), Washington, D.C.
PUB DATE	25 Jul 81
GRANT	NIE-G-78-0052
NOTE	32p.; Appendix A removed due to copyright restrictions.
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Cognitive Ability; *Cognitive Measurement; Discourse Analysis; *Educational Research; Elementary Education; *Listening Comprehension; *Memory; *Reading Comprehension; Word Recognition

ABSTRACT

To understand a text, a reader must engage in three important cognitive activities--recognition, comprehension, and memory. Based on this premise, two experiments were conducted with children to assess individual and developmental differences in speed of word recognition and how these differences related to performance on a variety of memory tasks. One unexpected finding was that although rapidity of word recognition increased sharply and continuously from grade one to grade nine, development was independent of this increase. An important implication of the two experiments is that developmental gains in rapidity of word recognition may have no necessary relation to memory improvement, but that individual differences may go hand in hand with differences among the same individuals in memory skills. Two other experiments with school-aged children revealed that certain effects studied extensively in sentence verification (the negation and comparator effects) replicated well in sentence completion, that stable differences among individual children did exist, and that the differences probably derived from variation in the efficiency with which different children executed elementary cognitive processes. (The full report of the first two experiments is appended.) (HOD)

FINAL REPORT

Children's Understanding of Written and Spoken Discourse

NIE-G78-0052

National Institute of Education

Alex Cherry Wilkinson
Principal Investigator
Psychology Department
University of Wisconsin
1202 W. Johnson Street
Madison, WI 53706

July 25, 1981

I. Introduction

"He correctly read all the words in the text, but he failed to understand it." Statements of this kind are often made by teachers of children in the early and middle grades.

If the child in question is in the early grades, some educators and psychologists would claim that the child is not really reading; he is merely saying words. Underlying this claim is an assumption that the child may have had sufficient intellectual prowess to understand the text, but was handicapped by having to read it. Perhaps the child would have comprehended better if someone else had read the text to him, and the child had simply listened.

In an older child, one who has reached the middle grades and is able to read words accurately, poor comprehension is typically attributed to difficulties different from those that may handicap a younger child. One might suspect that the older child would comprehend the text no better by listening than by reading. Perhaps the barrier to understanding for the older child is unrelated to the modality in which the message is expressed. Rather, the difficulty may reside in factors such as the complexity of the text, the reader's familiarity with the domain of ideas from which the text is drawn, and the child's skill in recognizing implicitly and explicitly stated meanings.

This contrast between younger and older readers is closely related to a distinction between developmental and individual differences in comprehension. The younger child's failure to read effectively for the meaning of a text may be a handicap specific to the mode of reading. Development of reading skills removes this handicap. The improved understanding of an older child is, therefore, partly a result of developmental processes involved in acquiring reading skills. Poor reading skills may not be the only impediment to understanding, however, and difficulties unrelated to basic reading skills may hinder novice readers and persist even in skilled readers. These difficulties may be viewed as individual differences. Some of the relevant dimensions of differences among individuals may be their background knowledge of the material to be understood and their facility in such fundamental cognitive processes as inference, storage of concepts in memory, and retrieval from memory. The distinction between developmental and individual difference is, to be sure, neither simple nor absolute. The two kinds of differences may coexist in a child. A poor reader in the middle grades may suffer deficits that are partly developmental, partly individual. Nevertheless, the distinction proves to be valuable as a device for understanding possible causes of children's difficulties in seeing meaning in what they read.

The general goal of the research described in this report was to examine systematically individual and developmental differences in cognitive abilities related to reading. The importance of research in this area is implied by investigators who have reviewed the state of our current knowledge of the psychology of reading. Gleitman and Rozin (1977) proposed that literacy should be defined as the ability to comprehend at least as well by reading as by listening. This proposal was accompanied, however, by the discouraging observation that too little is known about proper methods for comparing reading and listening comprehension, and that, as a consequence, there is no simple way to assess whether an individual may be considered literate under the proposed definition. Some of the complexities of assessment are discussed later in this report.

Significantly, the book containing the article by Gleitman and Rozin, although intended, apparently, to be a book on all psychologically important aspects of reading, contains several chapters on visual recognition of words but only one chapter on comprehension. A similar imbalance is evident in an otherwise comprehensive book, The Psychology of Reading, by Gibson and Levin (1975). The imbalance in both books is not the fault of the authors. As Gibson and Levin rightly observed (p. 392), "We know far less, in terms of basic research, at least, about factors that influence comprehension of sentences and longer passages of discourse in reading than we do about factors that influence recognition of individual words." Research in the past five years has begun to correct this imbalance, but it still exists as a problem.

The imbalance is unfortunate. Our knowledge of the psychology of comprehension is meager, yet comprehension is the essence of literacy, if attainment of literacy is taken to be the achievement of equal facility in understanding spoken and printed meanings. Moreover, an adequate theoretical and empirical account of comprehension is critical for understanding and alleviating the problems of poor readers, who often recognize printed words well enough but fail, nevertheless, to comprehend.

The remainder of this report addresses these issues in the following manner. First, a theory is described which outlines cognitive interrelations among processes of visual recognition, comprehension, and memory. The theory describes how these component processes interact during the act of reading for meaning--how the processes interact effectively in readers who understand well and how various limitations on the component processes may cause poor understanding. Second, results are reported for a series of five experiments conducted under the terms of grant NIE-G78-0052.

II. A Theory of Reading for Meaning

Reading to understand is surely a complex and elaborate cognitive skill. No complete theory of that skill is available. Several theorists have proposed models for facets of the larger process. There are models of word recognition (e.g., Massaro, 1975; Rumelhart & Siple, 1974), of parsing and remembering propositional statements (e.g., Anderson, 1976; Kintsch, 1977), and of recognizing highly abstract forms of meaning and textual structure (e.g., Simon & Hayes, 1976; Thorndyke, 1977).

The theory behind the present proposal is somewhat different from the preceding theories. Rather than present a detailed account of an isolated facet of reading comprehension, this report describes a general view of the holistic process of finding meaning in print. Discussion of the theory will be somewhat brief. A more thorough presentation is available, however, in Wilkinson (1980).

A. Essence of the Theory

To understand a text, a reader must engage in many cognitive activities. Three of the most important activities are processes of visual recognition, comprehension, and memory. The four words underlined in the preceding sentences are central to the theory. It seems best, therefore, to begin by defining these terms.

Recognition is identification of a printed configuration as a known item. The item is probably no larger than a short phrase (Rayner & McConkie, 1977). The perceptual process of recognition may, but need not, be accompanied by association of the recognized item with its sound, its meaning, or both.

Comprehension is the set of constructive and inferential processes (cf. Paris & Lindauer, 1977) by which an internal representation of the meaning of a text is derived. The representation may include propositional meanings (e.g., Anderson, 1976; Kintsch, 1977) as well as higher-order information about the structure of the text (Mandler & Johnson, 1977; Thorndyke, 1977). Effective comprehension is the construction of a representation which is internally cohesive, veridical, and well linked with the understander's prior knowledge (Greeno, 1977).

Memory is the set of processes that store and retrieve the outputs of recognition and comprehension.

Understanding, as the term will be used from now on, is to be distinguished from comprehension. Understanding is the holistic process that encompasses all of recognizing, comprehending, and remembering.

There is more than one way in which processes of recognition, comprehension, and memory might be organized in the cognitive activity of understanding. One possible form of organization assumes an interactive model of reading (Stanovich, 1980; Wilkinson, 1980). In an interactive model, processes of recognition, comprehension, and memory are all interdependent. Any of the three processes may hinder or facilitate the operations of any other. In addition, the processes are all assumed to draw upon a common pool of cognitive resources, although it is neither known nor hypothesized whether the common resources are central attentional capacity, limited buffer storage, shared informational pathways, or some other form of interdependency. One purpose of the proposed research is to obtain evidence that may reveal more clearly the nature of the interdependence among recognition, comprehension, and memory. That there is some form of shared limitation on the three processes seems likely, however.

B. Limitations on the Reader's Understanding

The preceding discussion suggests that there are several ways in which a reader might fail to understand a text. A taxonomy of causes for poor understanding must include the following.

Recognition limits: Restrictions on the use of external data. One kind of limitation may occur in a young child who has a small sight vocabulary or in a more skillful reader who scans a passage too quickly. Understanding of the passage may be impaired in either case by failure to recognize words or phrases that are critical to the meaning. This impairment is similar (but probably not identical) to what Norman and Bobrow (1975) called a signal data limit. Recognized words are the data on which comprehension depends, and the data secured from the printed page are so crude that an adequate understanding is not attained.

Comprehension limits: Restrictions on the availability of internal data. A limitation of another kind might arise when words are accurately recognized but an essential process of comprehension is not available to the reader. For example, one who cannot do mental arithmetic would fail when asked to verify, without the aid of pencil and paper, the statement, "Two hundred and thirty-four minus sixty-six is one hundred sixty-eight." Failure to comprehend the arithmetical statement may be defined operationally as inability to verify its correctness (Trabasso, 1972). One might fail to verify the statement despite accurate recognition of all the constituent words. The difficulty, then, is not that data from external sources are poor. On the contrary, the printed words are accurately recognized. Rather, the cause of failure is a paucity of internal data; veridical comprehension requires capabilities in mental arithmetic that the reader does not possess. Norman and Bobrow (1975) described cognitive phenomena of this sort and called them internal data limits.

Memory limits: Failure to store or retrieve what was comprehended. A third restriction may arise when the reader comprehends the important concepts but fails to save or recall them. The difficulty could be in the process of storage. An adequate representation of a text might be constructed in a short-term store but never properly transferred to a more permanent repository. Alternatively, there could be problems of retrieval. A properly stored representation may decay or become inaccessible with the passage of time and the intervention of new experiences.

Interactive limits: Competition for shared cognitive resources. Limitation of another kind may arise when the reader accurately recognizes the essential words, possesses the constructive and inferential operations necessary to comprehend the intended meaning, and is competent in storing and recalling ideas. In this case, there may be processing resources used jointly for recognition, comprehension, and memory. Competition for these resources may diminish the reader's ability to infer meanings or to store them in an easily retrievable form. If, for example, recognition is accomplished in an inefficient manner demanding a generous allocation of cognitive resources, only very limited mental capacities may remain to be deployed for comprehension and memory. It is as if so much effort were given to accurate recognition that little is left to be given to other processes. In the terminology of Norman and Bobrow (1975) the reader's performance is resource-limited in this case. Restricted understanding is not caused by forgetting or by inadequacy of internal or external data. Rather, the fault lies in competition for cognitive resources that cannot in the end meet all the demands placed upon them. LaBerge and Samuels (1974) have argued that an understanding of limited attentional resources is an important goal for research on reading.

C. Individual and Developmental Differences

What characteristics of a child may be expected to change in the usual course of development, and what characteristics are resistant to change? In a reader who understands poorly, what combination of developmental lags and individual qualities are responsible for the child's difficulties? How are these difficulties related to the several kinds of cognitive limitation discussed above? These questions are important but are not easily answered.

A beginning can be made by considering in a more precise manner the kinds of differences among children that could occur. Differences of one kind are stable variations among individuals. An example could be the efficiency and capacity of short-term memory. Individual differences in short-term memory appear to provide a partial explanation for variation in verbal ability among adults who are highly skilled readers (Hunt, Lunneborg, & Lewis, 1975). Perhaps such memory differences explain some portion of the variability in understanding at all levels of development. Thus memory limits, as defined above, could be a source of stable variation.

Experiential variations among individuals could also occur in understanding. For example, a reader's background knowledge of a topic may be viewed, in certain contexts, as a "frame" (Minsky, 1975) of relationships among concepts expected to be found in a text that is being read. The frames used by different individuals may vary in amount of detail, degree of suitability for the text at hand, and similar factors. With relevant experience, perhaps almost any individual could acquire an adequate frame for some domain of knowledge, such as a frame for understanding cooking recipes. As this example suggests, experiential variations may be similar to stable variations in being relatively unaffected by modal developmental processes, but dissimilar to them in being ameliorable by a specific program of instruction. Some cases of comprehension limits, as defined above, may be experiential.

Finally, developmental differences are variations among individuals which change with age and which change in similar ways over a broad range of environmental and experiential influences on the individual (Wohlwill, 1971). The development of deliberate approaches to memorization may be an example (Hagen, Jongeward, & Kail, 1975; Wagner, 1974). Specific experience may influence how speedily development occurs or how far it progresses, but not whether it occurs. Development of skill in managing one's own cognitive system could be viewed as a developmental difference in vulnerability to interactive limitation.

This set of distinctions seems relevant to the teaching of reading in two ways. First, it would be valuable to know which aspects of the understanding process may be altered by instruction and which may not. Second, assessment of an individual's standing on dimensions of developmental, experiential, and stable variation would seem to be a prerequisite for devising a program of instruction well suited to that individual.

III. Results of Experiments

In all, five experiments were performed. Four of them formed a longitudinal project in which a single sample of children was tested in a variety of ways from 1978 to 1980. The longitudinal project provided data concerning both individual and developmental differences among the children in the sample. In the fifth experiment, a sample of young adults was tested in 1981. The adult study and two of the studies with children examined the relation between rapidity in initially recognizing words and accuracy in later remembering them. The remaining two studies with children examined relations between comprehension and memory.

A. Recognition and Memory

There were three major reasons for conducting the experiments with children concerning the relation between rapid recognition of words and memory of the words. First, under the theoretical principles discussed above, a child's reading ability might be restricted by an interactive limitation in which the child is handicapped by inefficiency in coordinating processes of recogntion and memory. Second, there is little doubt that during the first few years of schooling, the speed with which children recognize words increases dramatically (Doehring, 1976; Friedrich, Schadler, & Juola, 1979), but there is, at the same time, little research on whether this increase has any affect on children's ability to remember. Finally, it is known that memory span correlates well with reading ability (e.g., Stevenson, Parker, Wilkinson, Hegion, & Fish, 1976), and it has been proposed that developmental improvement in memory span derives from increasing efficiency in initially recognizing the items that are to be remembered (Huttenlocher & Burke, 1976).

The general strategy of the two experiments with children was to assess individual and developmental differences in speed of word recognition and then to examine how these differences related to performance on a variety of memory tasks. A sample of 144 children was tested during the 1978-79 and 1979-80 school years. The children were in grades 3-8 during the first year and 4-9 during the second. A report of the experiment in the first year is in press in the Journal of Experimental Child Psychology and will be published in August of this year; galley proofs of this publication are attached as Appendix A. Data from the second experiment have been analyzed completely and are currently being written for submission to a scientific journal. Only the major findings of these experiments are summarized in this report.

One unexpected finding was that although rapidity of word recognition increases sharply and continuously from grade 3 to grade 9, memory development is independent of this increase. In the first experiment, it was found that memory variables such as storage and retrieval in free recall and memory for the items in a series and their order all exhibited a pattern of developmental growth different from the growth pattern for rapidity of initial visual recognition. For example, developmental gains occurred early but not late within the grade levels that were studied for a measure of storage in free recall, but growth was continuous over these grades for a measure of efficient word recognition. In the second experiment, several memory tasks were presented under two conditions, one that made it difficult for the children to initially recognize the words they were to remember and another condition that made word recognition easier. If younger children's poorer memory were partly the result of inefficiency in word recognition, their memory performance should have been handicapped more severely in the difficult condition than would be the case for older children. In fact, however, the handicap was equal and consistent across the broad range of ages in the experimental sample. Finally, correlations across the two experiments showed that there were stable individual differences among the children in such aspects of memory as rapid storage and efficient retrieval. There was some evidence, however, that these correlations did depend in part on individual differences in efficiency of word recognition.

Thus an important implication of the two experiments with children is that developmental gains in rapidity of word recognition may have no neccessary relation with memory improvement, but that individual differences in word recognition may go hand in hand with differences among the same individuals in memory skills. It remains an open question whether children who suffer difficulty in remembering what they read will benefit from training programs, such as courses in speed reading, that are designed primarily to accelerate word recognition.

Similar conclusions were implied by the experiment on the relation between word recognition and memory in adults. This experiment was motivated by research previously reported by Kolers (1975), Masson and Sala (1978), and Wilkinson, Guminski, Stanovich, and West (1981). The strategy of these various studies was to slow the speed with which adults readers could recognize the words of a sentence or text, in order to assess whether this handicap would affect how well they remembered the sentence or text. The previous studies by Kolers and by Masson and Sala showed that, surprisingly, one method of making reading difficult actually improved the reader's memory. The study by Wilkinson et al. showed that a different method of making reading difficult had no effect, either positive or negative, on a variety of memory variables. The study with adults that was performed under this grant was a direct comparison of the two methods. Data from the 30 adult subjects in this experiment have been completely analyzed, and they suggest that memory for the meaning of a text is protected from being damaged by slow word recognition if a certain condition obtains. A sufficient conditon for preventing memory loss is that immediately after reading a difficult sentence the reader must attend in some way to the meaning of that sentence. Under this conditon, memory for the sentence much later is unimpaired.

Plans are now underway to examine this condition systematically in an additional experiment. In the interim, two conlusions may be drawn tentatively. First, in adults as in children, visually reognizing words may be substantially independent of processes for remembering either the words themselves or the meanings they convey. Second, if the first conlusion is correct, then it can also be conluded that both word recognition and the interaction between word recognition and memory can be ruled out as significant causes of limited understanding in reading. These conclusions are tentative, however, because their relation to individual differences in cognitive abilities and in reading skill remains to be worked out.

B. Comprehension and Memory

In psychometric assessment of school-aged children, comprehension is often measured with tests that require completion of unfinished sentences. In psycholinguistic research usually done with adults, a common measure of comprehension is a sentence verification task in which a yes-no response is given to indicate whether a complete sentence is true or false. Are these two measures related? Are the cognitive processes identified by pscholinguistic research on sentence verification the same as those that are operative in psychometric assessment of sentence completion? Is the psychometric method correct in its tacit assumption that there are stable differences among children in the efficiency with which they perform the underlying cognitive processes? If stable differences exist, do they involve the interaction between cognitive processes of comprehending and remembering?

Two of the experiments that were performed under the grant examined these questions. The children in these experiments were 96 of the children from the longitudinal sample. They were tested twice, once in the spring of 1979 and then again in the fall of 1980. The two studies have been written up for publication and submitted to a scientific journal; the submitted manuscript is reproduced in Appendix B.

One goal of the experiments was to ascertain whether certain well established findings concerning sentence verification could be replicated in sentence completion. The results showed that certain effects studied extensively in sentence varification, the negation and comparator effects (Carpenter & Just, 1975; Trabasso, 1972; Wason, 1959), replicated well in sentence completion. In this respect, the experiments helped to build an empirical bridge between previous psycholinguistic research and traditional methods of assessing children's ability to comprehend.

A second goal of the experiments was to investigate the cognitive foundations of individual differences in ability on tests of sentence completion. In recent years, there have been numerous studies of individual differences in fundamental cognitive processes (see Hunt, 1978, and Sternberg, 1981 for reviews). To investigate whether there were differences in fundamental cognitive processes among the children in the longitudinal project, various measures of cognitive processing were derived from tests of sentence completion. For example, one measure was the magnitude of the negation effect for an individual child. Another example is that the children's efficiency in sentence completion was measured under conditions that imposed either an extra memory load or no extra load. These conditions made it possible to measure for each child the degree to which comprehension suffered when memory was heavily taxed. By correlating measures both within and across experiments, it was possible to assess whether individual differences were robust over time and across variations in experimental procedure. The correlations showed that stable differences among individual children did exist and that the differences probably derived from two underlying cognitive processes: the act of transforming the mental representation of an idea by negating it, and the process of managing demands on memory while concurrently trying to comprehend. This latter process of managing memory while comprehending is particularly interesting because it implies that individual differences may derive from interactive limitation concerning the coordination of memory and comprehension.

IV. References

Anderson, J. R. Language, memory, and thought. Hillsdale, NJ: Erlbaum, 1976.

Carpenter, P. A., & Just, M. A. Sentence comprehension: A psycholinguistic processing model of verification. Psychological Review, 1975, 82, 45-73.

Doehring, D. G. Acquisition of rapid reading responses. Monographs of the Society for Research in Child Development, 1976, 41 (2, Serial No. 165).

Friedrich, F. J., Schadler, M., & Juola, J. F. Developmental changes in units of processing in reading. Journal of Experimental Child Psychology, 1979, 29, 529-549.

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

Gleitman, L. R., & Rozin, P. The structure and acquisition of reading I: Relations between orthographies and the structure of language. In A. S. Reber and D. L. Scarborough (Eds.), Toward a psychology of reading. Hillsdale, NJ: Erlbaum, 1977.

Greeno, J. G. Process of understanding in problem solving. In N. J. Castellan, D. B. Pisoni, & G. R. Potts (Eds.), Cognitive theory (Vol. 2). Hillsdale, NJ: Erlbaum, 1977.

Hagen, J. W., Jongeward, R. H., & Kail, R. V. Cognitive perspectives on the development of memory. In H. W. Reese (Ed.), Advances in child development and behavior (Vol. 10). New York: Academic Press, 1975.

Hunt, E. B. Mechanics of verbal ability. Psychological Review, 1978, 85, 109-130.

Hunt, E., Lunneborg, C., & Lewis, J. What does it mean to be high verbal? Cognitive Psychology, 1975, 7, 194-227.

Huttenlocher, J., & Burke, D. Why does memory span increase with age? Cognitive Psychology, 1976, 8, 1-31.

Kintsch, W. Reading comprehension as a function of text structure. In A. S. Reber & D. L. Scarborough (Eds.), Toward a psychology of reading. Hillsdale, NJ: Erlbaum, 1977.

Kolers, P. A. Specificity of operations in sentence recognition. Cognitive Psychology, 1975, 7, 289-306.

LaBerge, D., & Samuels, S. J. Toward a theory of automatic information processing in reading. Cognitive Psychology, 1974, 6, 301-334.

Mandler, J. M., & Johnson, N. S. Remembrance of things parsed: Story structure and recall. Cognitive Psychology, 1977, 9, 111-153.

Massaro, D. W. Primary and secondary recognition in reading. In D. W. Massaro (Ed.), Understanding language: An information processing analysis of speech perception, reading, and psycholinguistics. New York: Academic Press, 1975.

Masson, M. E. J., & Sala, L. S. Interactive processes in sentence comprehension and recognition. Cognitive Psychology, 1978, 10, 244-279.

Minsky, M. A framework for representing knowledge. In P. H. Winston (Ed.), The psychology of computer vision. New York: McGraw-Hill, 1975.

Norman, D. A., & Bobrow, D. G. On data-limited and resource-limited processes. Cognitive Psychology, 1975, 7, 44-64.

Paris, S. G., & Lindauer, B. K. Constructive aspects of children's comprehension and memory. In R. V. Kail & J. W. Hagen (Eds.), Perspectives on the development of memory and cognition. Hillsdale, NJ: Erlbaum, 1977.

Rayner, K., & McConkie, G. W. Perceptual processes in reading: The perceptual spans. In A. S. Reber and D. L. Scarborough (Eds.), Toward a psychology of reading. Hillsdale, NJ: Erlbaum, 1977.

Rumelhart, D. E., & Siple, P. Process of recognizing tachistoscopically presented words. Psychological Review, 1974, 81, 99-118.

Simon, H. A., & Hayes, J. R. The understanding process: Problem isomorphs. Cognitive Psychology, 1976, 8, 165-190.

Stanovich, K. E. Toward an interactive-compensatory model of individual differences in reading ability. Reading Research Quarterly, 1980, 16, 32-71.

Sternberg, R. J. Intelligence and nonentrenchment. Journal of Educational Psychology, 1981, 73, 1-16.

Thorndyke, P. W. Cognitive structures in comprehension and memory of narrative discourse. Cognitive Psychology, 1977, 9, 77-110.

Trabasso, T. Mental operations in language comprehension. In J. B. Carroll & R. O. Freedle (Eds.), Language comprehension and the acquisition of knowledge. Washington, DC: Winston, 1972.

Wagner, D. A. The development of short-term and incidental memory: A cross-cultural study. Child Development, 1974, 45, 389-396.

Wason, P. C. The processing of positive and negative information. Quarterly Journal of Experimental Psychology, 1959, 11, 92-107.

Wilkinson, A. C. Children's understanding in reading and listening. Journal of Educational Psychology, 1980, 72, 561-574.

Wilkinson, A. C., Guminski, M., Stanovich, K. E., & West, R. F. Variable interaction between visual recognition and memory in oral reading. Journal of Experimental Psychology: Human Learning and Memory, 1981, 7, 111-119.

Wohwill, J. F. The age variable in psychological research. Psychological Review, 1970, 77, 49-64.

Appendix A

JOURNAL OF EXPERIMENTAL CHILD PSYCHOLOGY [redacted] (1981)

Growth Functions for Rapid Remembering **

ALEX CHERRY WILKINSON

University of Wisconsin-Madison

Children between ages 9 and 13 were tested for recognizing and remembering words from 6- and 12-word lists. Opportunities for using deliberate mnemonics were severely restricted. Developmental functions showed different growth patterns for remembering the items in a short list than for remembering order, and different patterns for storing items from a long list than for retrieving them. However, none of these functions was parallel to the growth function of rapid word recognition. This absence of parallel growth contradicted a hypothesis that memory develops when item recognition develops. The data suggested, instead, that modest but reliable gains in rapid processes of storage and retrieval contribute to memory development during middle childhood.

**The above article removed due to copyright restrictions.

Appendix B

Memory and Individual Differences in Children's Comprehension of Incomplete Sentences

Alex Cherry Wilkinson

Ronald Koestler

University of Wisconsin -- Madison

Running head: Sentence Comprehension

Tim Schultz assisted in collecting and analyzing the data of Experiment 1 and wrote an undergraduate honors thesis on the results. Susan Riley and Margaret DeMarinis also gave able assistance. The cooperation of the teachers and principals of the Oregon School District was invaluable. Funds for the research came from a grant from the National Institute of Education.

Requests for reprints should be sent to Alex Cherry Wilkinson, Psychology Department, University of Wisconsin, 1202 W. Johnson St., Madison, WI 53706.

Abstract

In two experiments on comprehension, a new paradigm was used in which a child first received an input specifying a serial order, and then gave a forced-choice response to complete a sentence testing comprehension of a vebal description of the order. The new paradigm provided an empirical bridge between a standard psychometric method of assessing comprehension and the method of sentence verification often used in psycholinguistic research. Two major effects known to occur in sentence verification, the negation and comparator effects, were replicated with the new paradigm of sentence completion, although differences between the paradigms were also suggested. How children selected a response for an incomplete sentence appeared to depend on how the input order was represented in memory. Individual differences in ability were examined by testing the same sample of children (10-14 years old) in both experiments. Correlations across and within experiments suggested that differences in ability were stable and derived from variance in two cognitive skills: executing a process of negation and managing memory loads.

Memory and Individual Differences in Children's Comprehension of Incomplete Sentences

In psychometric assessment of school-aged children, comprehension is often measured with tests that require completion of unfinished sentences. In psycholinguistic research usually done with adults, a common measure of comprehension is a sentence verification task in which a yes-no response is given to indicate whether a complete sentence is true or false. Are these two measures related? Are the cognitive processes identified by pscholinguistic research on sentence verification the same as those that are operative in psychometric assessment of sentence completion? Is the psychometric method correct in its tacit assumption that there are stable differences among children in the efficiency with which they perform the underlying cognitive processes? The present research examined these questions.

One purpose of our research was to ascertain whether certain well established findings concerning sentence verification could be replicated in sentence completion. Perhaps the oldest and most robust finding concerning sentence verification is the negation effect (Carpenter & Just, 1975; Trabasso, 1972; Wason, 1959). When the sentence to be verified is negative, it takes longer to be verified and is more likely to be verified erroneously than when the sentence is affirmative. This effect of negation holds whether the sentence to be verified is true or false. Theoretical accounts generally attribute the negation effect to a stage of the verification process that must occur for a negated sentence but is omitted for an affirmative one. We investigated whether a similar negation effect occurs in sentence completion.

Another effect often reported in studies of sentence verification concerns comparators. A good example is the pair of comparators above and below (Clark & Chase, 1972). Suppose one sees a picture of a * beneath which is a +. Two logically equivalent statements concerning this picture are The star is above the plus and The plus is below the star. The sentence with before is verified less rapidly and with more errors than is the sentence with above. The reason for this difference between comparators may be that above is held in memory in a semantically primitive, unmarked form, whereas below is stored as a marked derivative of above (Clark & Chase, 1972; Carpenter & Just, 1975). The more primitive meaning is presumed to be accessible more rapidly and with less likelihood of error. For our purposes, the important point of this example is that semantically related comparators may have different representations in memory and may differ, therefore, in their effect on comprehension. We investigated this possibility with two pairs of comparators: before-after and faster-slower.

When negation and comparator effects are considered together, it is important to ask whether they interact. What kind of interaction may exist and how it should be interpreted are unsettled and hotly debated questions (Carpenter & Just, 1975; Catlin & Jones, 1976; Shoben, 1978). The debate on this issue has neglected, we think, an important fact: An interaction between negation and comparator specifies an effect of position. Consider, for example, the following sentences, which are like the sentences we used in one of our experiments:

John was faster than Susan, and she was faster than Tom. (1)
Susan was faster than *** . (2)
Susan was not slower than *** . (3)

Sentences 2 and 3 are logically equivalent, and they can both be completed correctly with information from Sentence 1. In both cases, the correct response is the third name mentioned in Sentence 1; thus position three is correct. Significantly, if there were a crossover interaction between negation and comparator, then Sentences 2 and 3 would be statistically equivalent. The statistical interaction could mean that psychologically there is an effect of position, faster being like not slower because they both specify position three and, similarly, slower being like not faster because they both specify position one.

The possibility of a position effect in remembering and comprehending sentences like 1 is of interest because position effects are known to occur in cognitive processing of information that, unlike a sentence, has no syntactic structure. Most notably, position effects are known to occur in recalling items from a serially presented list and in comparing terms from an overlearned ordinal set (Potts, Banks, Kosslyn, Moyer, Riley, & Smith, 1978; Trabasso, 1977). Identifying position effects in sentence completion and specifying the conditions under which they occur could help to pinpoint fundamental properties of memorial representation that affect cognitive processing of both syntactically structured sentences and ordinally structured series. Consequently, our experiments examined whether position effects occur when the information needed to complete a sentence must be found in the memorial representation of a sentence or an ordered series.

Finally, our research was designed to investigate the cognitive foundations of individual differences in ability on tests of sentence completion. In recent years, there have been numerous studies of individual differences in fundamental cognitive processes (see Hunt, 1978, and Sternberg, 1979, for reviews). Some studies have shown that there are individual differences in the strategies that individuals prefer (e.g., MacLeod, Hunt, & Mathews, 1978), whereas others have suggested that individuals may use the same strategy with differing levels of efficiency (e.g., Keating & Bobbit, 1978). We assumed that there were differences in ability among the children we studied, and we assessed them in two experiments with various measures of sentence completion. For example, one measure was the magnitude of the negation effect for an individual child. By correlating measures both within and across experiments, we sought to establish whether individual differences were robust over time and across variations in experimental procedure.

Before turning to the details of our experiments, we wish to make explicit the scope of our goals. First, regarding individual differences, our primary goals were to establish the robustness of differences in ability and to find possible causes of the differences, rather than to identify the causes definitively. Thus, for example, the correlations we report may derive either from differences in the effectiveness of children's various strategies or from differences in how efficiently they executed a common strategy. Second, regarding sentence completion, our principal concern was to establish whether major empirical effects from the paradigm of sentence verification, especially the negation and comparator effects, were replicable in the new paradigm of sentence completion. Explicating a theoretical integration of the two paradigms is an important but separate task not undertaken in this article.

Method

In each of two experiments, we presented children with an input order that specified a three-term series and then tested the children's comprehension by having them complete a sentence concerning the input order. In both experiments, we manipulated the conditions of input in ways that were intended to reveal the role of memorial processes in sentence completion. Manipulation of input condition in the first experiment involved presenting inputs that formed a meaningful spoken or printed sentence, a meaningless printed word list, or a familiar sequence of printed letters. We expected that differences in the content of these inputs might result in different ways of representing the critical ordinal information in memory, with differential consequences for sentence completion. In the second experiment, input condition was manipulated with a dual-task technique. During input, the children were given, first, a sentence to be remembered and, second, an input order in a meaningful printed sentence. They were then tested in one of two ways. They either recalled the first input sentence or completed a sentence involving ordinal information from the second input sentence. Not knowing whether they would be tested for recall or order on any given trial, the children had to be prepared for both kinds of test. Their performance in this dual-task condition was compared with their performance in a control condition in which there was no extra sentence to be remembered. We reasoned that imposition of an added memory load in the dual-task condition might reveal how memorial processes are managed during sentence completion.

Subjects

In Experiment 1, the subjects were 96 children, aged 9-13 years. The subjects in Experiment 2 were 90 of these same children and 6 replacements. In both experiments, the sample of children included 12 boys and 12 girls at each of four consecutive school grades. The grades were 4-7 in the spring when Experiment 1 was conducted, and 5-8 in the following fall when Experiment 2 was conducted.

Experiment 1

Materials. The test item for an individual trial was a white card printed on both sides. One side contained an input order, and the other, a test sentence. Both sides were printed in large plain letters with a 6-pitch typewriter.

The input order was in upper case and was one of three types: (a) three consecutive letters from the alphabet that formed an overlearned triple (ABC, EFG, TUV, or XYZ); (b) a meaningful three-word sentence structured grammatically as SVO (subject-verb-object, e.g., KIDS PLAY GAMES); or (c) a meaningless string of three nouns or three verbs (e.g., CARS QUEENS KIDS). Twelve three-word sentences were prepared; then the three-word strings were made by randomly selecting either nouns or verbs from the sentences, subject to the constraint that no two words from the same sentence were selected for the same string. Two additional sentences and strings were made for practice.

The test sentence was either a choice sentence or a control sentence. Symbolically, a choice sentence, as printed had one of the two general forms:

$$2\ V \begin{matrix} 1 \\ 3 \end{matrix} \quad \text{or} \quad 2\ V \begin{matrix} 3 \\ 1 \end{matrix}$$

Similarly, a control sentence had either of the following forms:

2 V 1 or 2 V 3.

In these symbolic forms, a number stands for the element with the corresponding ordinal position in the input order, and V stands for a verb phrase. The verb phrase was one of the four that can be generated by combining was or was not with before or after. A child saw test sentences with actual words, not the symbolic forms. An example is

KIDS.
PLAY was before
GAMES.

Capitalization of the test sentences was as in the example. In the case of choice sentences, the word that completed the sentence correctly was equally often on the top or the bottom. In the case of control sentences, the sentence was always correct as printed.

An audio tape was made for the sentence inputs and their associated test sentences. On the tape, a male or female speaker said the input sentence, then said all but the last word of the test sentence.

Procedure. A child had 2 practice trials and 12 test trials in each of four conditions: (a) read sentences: The child read input and test sentences. (b) look-listen to sentences: The input and test sentences were presented by audio tape through headphones with the corresponding printed sentences concurrently available to be seen on cards. The child said only the word that completed the test sentence. (c) read words: The child read word-string inputs and test sentences. (d) read letters: The child read input letter-triples and test sentences. The 12 test trials in a given condition formed a 4 x 3 design. The four possible verb phrases in the test sentence were crossed with three kinds of sentence completion: top choice correct, bottom choice correct, and control with one choice.

A single set of the 4 letter-triples, 12 input sentences, and 12 word-strings was permuted to generate six counterbalanced arrangements. These arrangements balanced type of test sentence associated with a given input, presentation order of the input types, and identity of the experimenter. Two arrangements were assigned to each of three experimenters, each of whom tested one third of the subjects at each age.

Subjects were tested individually in a single session lasting 20-30 minutes. The data for each trial were times recorded by the experimenter with a stopwatch and, in the case of choice sentences, the correctness of the subject's choice. Times recorded by the experimenter were the choice time from the beginning of a choice sentence until the child said the word that completed it, and the analogous control time on a control sentence. Choice times were averaged over the two trials (top-correct and bottom-correct) for each combination of input condition and verb phrase of the test sentence. Then the single corresponding control time was subtracted from this average to yield an estimate of decision time.

Experiment 2

The second experiment was presented on a CRT monitor under the control of a microcomputer. Reaction times were measured by a clock in the computer with an accuracy of ± 8 msec.

Materials. Some trials had a memory sentence. All trials had both an order sentence and a response segment.

1. A memory sentence began with a temporal clause in the past tense, consisting of one of the possible combinations of a number from the set (two, four, six) and a temporal unit from the set (hours, days, weeks, months). The temporal clause was followed by a common surname prefixed by Mr. or Mrs. A stopwatch clause followed. It was the invariant phrase "timed the children with a stopwatch to find out who could." The last portion of the sentence, the activity phrase, named a familiar activity and ended with the words "the fastest." An example is: "Six days ago, Mrs. Young timed the

children with a stopwatch to find out who could tie their shoes the fastest."

2. Order sentences had an invariant grammatical structure but differed in proper names and pronouns. Names were arranged in sets of three, either two boys and one girl, or two girls and one boy, where the odd sexed name referred to as the pivot, occurred in the middle of the sentence, between the two referents. An example is: "Tom was faster than Judy, and she was faster than Jim."

3. A response segment was one of two types: (a) On half of the trials that included a memory sentence, the word "stopwatch" was presented after the order sentence as a recall cue. (b) On all other trials, the order sentence was followed by a test sentence in one of the four forms: (pivot) was (faster, slower, not faster, not slower) than *** . Following this phrase the subject was shown both referents or only the referent that correctly completed the test sentence. When both referents were shown, they were on the same line and the correct referent was equally likely to be on the left or the right.

Procedure. Each subject received 12 trials in each of three conditions. (a) order only: Subjects were shown an order sentence, a test sentence, and one or two referents. On choice trials with two referents, the subject responded by pressing one of two buttons, which were held continuously in the left and right hands. On control trials with one referent, subjects were instructed to press either button as quickly as possible. (b) dual-order: Subjects were shown a memory sentence, an order sentence, a test sentence, and one or two referents. The procedure for responding was the same as in the order-only condition. (c) dual-memory: Subjects were shown a memory sentence, an order sentence, and the "stopwatch" cue. On seeing the cue, subjects were to try to recall the memory sentence verbatim.

In each condition the subjects were required to read every sentence aloud at a normal and comfortable rate. During a trial, as soon as the subject finished reading a sentence, the experimenter pressed a key on the computer keyboard to remove the sentence from the screen and replace it with the next sentence.

The three conditions were presented in two blocks; the order of these blocks was counterbalanced across subjects. One block contained the order-only trials. The other block contained dual-memory and dual-order trials, randomly intermixed; the randomization was done independently for each subject. In this block, subjects did not know whether they would be tested for memory or order until the response segment of the trial began.

Subjects were tested individually in a single session lasting 20-30 minutes by one of two experimenters. The data for each order-only and dual-order trial were reaction times measured by the computer, and in the case of choice trials, the correctness of the subject's choice. As in the first experiment, two times were of interest: the choice time from the onset of two referents to the subject's response, and the control time from the onset of a single correct referent to the pressing of either response button. Decision time was computed by averaging choice times for the left-correct and right-correct trials for one of the four kinds of test sentences, and then subtracting the control time for that same kind of test sentence. For each memory trial the subjects were given a memory score comprised of one point for each of the following items correctly recalled: the number in the temporal clause, the temporal unit in the temporal clause, the name, and the activity part of the activity phrase.

Statistical Analysis

With each child having two choice trials and one control trial for a given type of test sentence, the data were too sparse to provide reliable estimates of time and accuracy for every individual on every type of test sentence. Thus it was advisable to average. For analyses concerned with memory processes and developmental differences, averages were computed for a sentence type across subgroups of children. For analyses concerned with individual differences, averages were computed for a child across susets of sentences.

In computing these averages, two corrective steps were taken. First, mean decision times were computed as trimmed means, excluding a fixed number of both high and low values. According to Mosteller and Tukey (1977), trimmed means are more efficient than ordinary means for distributions which have markedly long tails, as did those of the decision times in the present experiments. Second, to prevent restriction of variance for proportions that sometimes approached 1.0, proportion correct was transformed to logistic accuracy, as recommended by Mosteller and Tukey (1977). The formula for this transformation is $y = \ln(p / (1 - p))$, where for x correct out of N, $p = (x + .167) / (N + .333)$.

In analyses concerning memory processes and developmental differences, logistic accuracy with $N = 12$ and mean decision time with N trimmed to 8 were computed across the 12 children in a cell of the experimental design. These means were then used as the data for analyses of covariance. Two simple analyses of covariance were done, one on accuracy with time as the covariate and the other on time with accuracy as the covariate. Factors in these analyses were age, sex, test sentence, input condition, and all two-way interactions among these factors. Higher interactions were pooled as error. These analyses made the simplifying assumption of a single common slope for the covariate. Of greater interest was another analysis, called the full covariance analysis, in which the slope of the covariate was allowed to vary according to age, sex, test sentence, input condition, and the two-way interactions. In this analysis, accuracy was the dependent variable and time the covariate because the principal concern was to examine accuracy as a function of the time taken to decide on a response. Effects that were empirically null in the full covariance analysis were successively deleted from the full covariance analysis until only significant effects remained. In all analyses, age was coded as the contrast between children in the two older groups combined (averaging 13 years) and those in the two younger groups (averaging 11 years). Comparable analyses using the linear trend in school grade yielded comparable results but will not be reported.

To analyze individual differences, logistic accuracy and trimmed mean decision time were computed for groups of test sentences that had been found in the covariance analyses to be appropriately grouped together. Values were computed for each child individually in each experiment. All the child's data were used in computing logistic accuracy for a particular group of sentences. The slowest and fastest times for that child on a group of sentences were trimmed from that child's mean decision time. Only the data of children who were subjects in both experiments were used in the analyses of individual differences.

Results

Experiment 1: Analyses Concerning Memory and Development

A simple covariance analysis showed that decision time, adjusted for logistic accuracy, was faster for looking and listening to sentence input (.91 sec) than for reading sentence input (1.28 sec), $F(1, 102) = 7.75$, $p < .01$. This finding probably indicates that in looking-listening, the child could begin to decide on a response to the test sentence as the test was presented and so could respond quickly. In reading, however, the requirement to say the test sentence aloud may have inhibited the decision process until oral reading had ended.

A second effect in the same analysis was that adjusted maximum decision time was faster for printed and spoken sentences (1.10 sec) than for word-lists and letter-triples (1.25 sec), $F(1, 102) = 6.10$, $p < .05$. This finding suggests that during decision making, access to the memorial representation of a sentence was faster than to that of a series.

In one of the simple covariance analyses, the slope for logistic accuracy was virtually zero, and so was the slope for decision time in the other simple analysis, both F's < 1. However, the full covariance analysis showed that the slope for decision time in the time-accuracy function varied according to an interaction between age and test sentence, $F(3, 112) = 3.11$, $p < .05$. Inspection of the slopes revealed that in all but one case the slope was virtually zero. The single exception was when older children responded to test sentences containing <u>before</u>. Consequently, a model was fit to the data in which this slope alone was free to differ from zero. The estimate of this slope was .98 units of logistic accuracy per sec, which is approximately a gain of .20 per sec in proportion correct.

Given one nonzero slope, it was necessary to select a benchmark time at which mean accuracies could be compared. The time chosen for this purpose was 1 sec, which was approximately the median time and was within the range of observed times at each of the four ages investigated in the study. Estimated accuracy at 1 sec, as computed from the model with one nonzero slope, will be called <u>benchmark accuracy</u>.

Converting from logistic values back to proportion correct, the benchmark accuracy of older children (.75) was greater than that of younger children (.66), $F(1, 119) = 13.53$, $p < .001$. In addition, there was a comparator effect, <u>after</u> producing greater benchmark accuracy (.76) than <u>before</u> (.63), $F(1, 119) = 9.65$, $p < .01$. These mean differences concerning age and preposition must be interpreted with an eye to the previously mentioned effects concerning slopes. Within the range of observed decision times, the averaged time-accuracy functions for younger children were two parallel lines, both with zero slope, the <u>after</u> line being .13 higher than the <u>before</u> line. The zero slopes may mean that the younger children responded at times when their accuracy had reached an asymptotic plateau. In contrast, the fitted functions for older children were, first, a line with zero slope but high benchmark accuracy for <u>after</u> and, second, a line with positive slope but with benchmark accuracy lower by .13 for <u>before</u>. Extrapolation implied that the two lines for the older children would converge at 1.6 sec, but this projection is tenuous because it is at the boundary of the observed decision times. These results imply an interesting kind of development. Younger children were handicapped by <u>before</u>, as were the older children, but the older ones were able to compensate for this handicap, at least in part, if they delayed their response.

In addition to this developmental interaction, effects were found concerning memory processes that, within the age range and statistical power of the experiment, were developmentally constant. Two such effects were interactions between input condition and an aspect of the test sentence. First, the preposition effect was in one direction for familiar letter-triples and in the opposite direction for unfamiliar word-lists, $F(1, 119) = 7.90$, $p < .01$. Means for this interaction are in Table 1. Second, letters and words showed opposite position effects, $F(1, 119) = 13.67$, $p < .001$. The relevant means are in Table 2. Note that the tabled means show no position effect for sentences and a preposition effect for sentences resembling that for letters.

Insert Tables 1 and 2 here.

One explanation of these results is that there were directional biases deriving from the manner in which various inputs were represented and accessed in memory. With familiar letter-triples, a bias favoring both after and the first position could have resulted from a serial representation. Perhaps the letter inputs were available in long-term memory as serially ordered sets most readily accessed at their initial element. In constrast, the unfamiliar word lists may have been held in a temporary representation for which, as in free recall, recent elements were likely to be accessed first for output. Thus a bias favoring both before and the third position may have resulted from factors similar to those responsible for the well known recency effect. More problematic is the finding that sentences, unlike the other inputs, showed no position effect. A plausible argument can be made that sentences produced different results because their memorial representation was different from that of both letter-triples and word-lists. Rather than being stored as an ordered series, a sentence may have been represented as a hierarchically organized proposition. Perhaps access to the proposition was gained through the verb, which in a simple SVO construction could well have been equidistant in memory from the two response choices. Kintsch's (1974) theory of propositional representations in memory seems compatible with this account, because the verb , according to Kintsch, is the leading constituent in the memorial representation of an SVO construction. Admittedly, the present data do not establish conclusively how the various inputs were memorially represented. Without some assumption of differing representation, however, it would be difficult to explain the observed directional biases.

Finally, a strong effect on benchmark accuracy was observed for negation, $F(1,119) = 7.53$, $p < .01$. Children were more accurate for affirmative test sentences (.75) than for negatives (.64). No interaction compromised this effect.

Experiment 2: Analyses Concerning Memory and Development

In a dual-task paradigm like that of the second experiment, the secondary task should have certain properties. To assure that it requires attention, the secondary task should be difficult but not overly so, avoiding both floor and ceiling effects. In addition, it should be equally difficult for different groups of children; otherwise, group differences on the primary task might be unwanted side effects of differences in ability to manage the secondary task. The recall task had the desired properties. Overall, the mean proportion of items correctly recalled was .52. An analysis of variance showed no hint of age or sex differences, all Fs < 1. There were, however, individual differences among children in excess of the variation in measured abilities that would be expected if all children had the same true ability. Under an assumption of binomial sampling and equal true ability, the expected standard deviation of measured abilities on the secondary recall task would be .072; empirically it was .144. The large variation among individuals raises the possibility that different children had different strategies for managing the secondary recall task, some strategies being more effective than others. Nevertheless, the task was appropriately difficult on the average, and differences in strategy, if present, left group averages virtually equal across age and sex.

A full covariance analysis showed that imposition of the secondary recall task penalized performance on the primary task of sentence completion. The mean proportion correct on the primary task was .62 with the secondary task and .68 without it; $F(1, 116) = 8.64$, $p < .01$. This result implies that the primary task of sentence completion made demands on memory that conflicted with those of the secondary recall task. Importantly, there was no indication that the penalty of the secondary task interacted with any other factor, not even marginally. Thus, for example, any developmental differences on the primary task were, in the present data, a consequence of factors other than developmental improvement in ability to manage competing demands on memory.

The means reported in the preceding paragraph, and others to be reported below, were computed by untransforming logistic values from the full covariance analysis. The untransformed values estimate accuracy at a benchmark decision time of 250 msec. This benchmark value was, like that of the first study, approximately the median time and was within the range of observed times at all ages. The 250 msec benchmark of Experiment 2 is markedly shorter than the 1 sec benchmark of Experiment 1, however. This difference may derive in part from the use of a stopwatch in the first experiment and a computer in the second. In addition, under the procedure of the first study, times on control trials could have been unreasonably short because the subject knew while reading the initial part of the test sentence on a control trial that there was only one choice for response. Consequently, the control time may have been speeded and, when subtracted from the choice time, it could have produced a generous estimate of decision time. In contrast, when a child read the initial part of a test sentence in the second study, there was no way to know whether two choices would appear for response or only one. There was probably no speed-up of the control time, therefore, and the estimate of decision time was correspondingly more conservative.

Benchmark times are reported for Experiment 2 because, as in the earlier experiment, the slopes varied. Unexpectedly, they varied by sex. The slope of the time-accuracy function was -.33 per sec in logistic units for girls, and +.99 per sec for boys, $F(1, 116) = 6.45$, $p < .02$. Translating to proportion correct, these slopes mean that over the interval from 0 to 500 msec decision time, girls' accuracy would decline slightly and nonsignificantly from .67 to .63, whereas boys' accuracy would rise reliably from .59 to .71. The constancy of the girls' accuracy is sensible if, in fact, the girls rarely responded with decision times near zero, preferring to wait until they had reached a point at which no further gain in accuracy could be achieved. The boys, however, may have preferred a strategy emphasizing speed, resulting in some quick but inaccurate responses and in better accuracy when responese were delayed. Data from a simple covariance analysis of decision times, adjusted for logistic accuracy, supported this interpretation of the sex difference. Boys' times averaged 49 msec faster than girls', $F(1, 108) = 6.71$, $p = .01$, a speed-up of about 20% relative to the benchmark decision time of 250 msec.

Developmentally, three effects involving age were found in the full covariance analysis. First, the benchmark accuracy of older children (.68) was greater than that of younger children (.63), $F(1, 116) = 5.52$, $p < .05$. This difference was present, however, only when the correct choice corresponded to the first position in the order sentence. As shown in Table 3, age interacted with Position, $F(1, 116) = 7.73$, $p < .01$. A satisfying explanation of this interaction is elusive. There was also a significant interaction between age and adjective, as shown in Table 4, $F(1, 116) = 5.95$, $p < .05$. Although statistically significant, this interaction seems artificial because a floor effect attenuated the age difference for the adjective slower.

Insert Tables 3 and 4 here.

Notably, neither of the interactions in Tables 3 and 4 is a crossover interaction; nor does either have a compelling theoretical interpretation. It is reasonable, therefore, to examine the corresponding main effects. The modest effect of position, shown in the margin of Table 3, was significant, $F(1, 116) = 3.88$, $p < .05$. In direction, this effect is like the position effect for word-lists in Experiment 1 (see Table [illegible]). It may be that a bias favoring position 3 occurs whenever the input order is arbitrary, for both the word-lists in Experiment 1 and the sequences of names in the order sentences in Experiment 2 were haphazard. In magnitude, however, the position effect was so small in the present experiment that it barely differed from the null effect for sentence inputs in the earlier experiment. Thus one might argue that whenever the input order is embedded in a sentence, the position effect will be negligible. Two interpretations may be given, therefore, concerning the relation between sentence content and positional bias. Only additional data can establish which of the two is better.

In contrast to the ambiguous results concerning position effects, the main effect of adjective, shown in the margin of Table 3, is convincingly large, $F(1, 116) = 42.51$, $p < .001$. In magnitude, this comparator effect is comparable to the preposition effects in Experiment 1 (see Table 1). Two factors may have contributed to the adjective effect. First, the adjective in the order sentence was faster; when this adjective also appeared in the test sentence, it may have cued retrieval of the order sentence. Second, when faster was in the test sentence, the correct response was the referent that had followed this adjective in the order sentence; thus the ordinal structures of the two sentences were compatible. But is ordinal compatibility always helpful? There was ordinal compatibility for test sentences with before in the previous experiment, and, indeed, performance was generally better on these sentences. The opposite was true, however, when the input was an overlearned letter-triple. Perhaps what mattered was not compatibility between temporal order of input and sequence of mention in the test sentence. Instead, the critical factor may have been a tacit assumption that as elements of the input order were retrieved from memory, the one most readily retrieved was likely to be the correct one. This account squares with the findings from both experiments.

Finally, benchmark accuracy was greater for affirmative test sentences (.68) than for negatives (.63), $F(1, 116) = 5.43$, $p < .05$. This effect was somewhat smaller than in Experiment 1, although not unreasonably so. The effect could have been attenuated by the generally lower level of performance in the second experiment, in which mean benchmark accuracy was .65 as compared to .70 in the first experiment.

Both Experiments: Analyses Concerning Individual Differences

To investigate the stability of individual differences among children, averages were computed across test sentences for each child. For example, from the data of Experiment 1 four values were computed: (a) the logistic transform of mean accuracy over all test sentences, (b) mean decision time trimmed to exclude the child's fastest and slowest responses, (c) the child's negation effect for accuracy, computed as the difference between logistic accuracy on affirmative and negative test sentences, and (d) the child's negation effect for decision time, computed as the difference between trimmed mean times on affirmative and negative tests. (Estimates of a child's individual position and preposition effects were not computed for this experiment, however, because these effects interacted with input condition, as reported in Tables 1 and 2). Variables (a) and (b) were used together as a composite indicator of the child's overall performance in the experiment, and variables (c) and (d) were used together as a composite indicator of the negation effect for that child. Similarly, accuracy and time variables computed from the data of Experiment 2 provided composite indicators of a child's overall performance, negation effect, adjective effect, position effect, and dual-task penalty. A final variable was the child's total score on the secondary recall task in Experiment 2.

Canonical correlations were computed among these composite indicators. A canonical correlation measures the association between two sets of variables; in the special case where one set contains one variable (i.e., the recall score) and the second set contains more than one, a canonical correlation is identical to a multiple correlation. Actually, the canonical correlations were computed twice, first on the original variables as defined above and then on residual variables computed by removing linear and quadratic trends in school grade from every original variable. The original and residual canonical correlations were nearly identical. Differences in absolute magnitude between the two correlations ranged from zero to .054 and averaged .016. Thus the correlations remained the same when developmental variance was removed.

Insert Table 5 here.

Table 5 shows the canonical correlations for the original variables. It should be noted that these variables are based on, at best, a modest number of trials per child and that most of the variables are difference scores, which are notorious for having low reliability. Thus the correlations may be attenuated. Despite this possibility, several correlations were significant. The correlation between overall performance in the two experiments indicates that the cognitive abilities underlying sentence completion are moderately stable over time (approximately nine months) and over nontrivial modifications of experimental design.

Notably, the negation effect was associated with overall performance. Of the four correlations between two estimates of the negation effect and two of overall performance, three were significant. Inspection of the canonical coefficients (analogous to regression coefficients) showed that children with higher overall performance had smaller deficits on negative sentences.

Finally two results concerning the dual-task penalty were interesting. First, the penalty was smaller for children with high overall performance than for those who were low overall. Recall that the dual-task penalty was not associated with age. It was, however, associated with individual differences in sentence completion. Apparently, ability to manage an extra memory-load varied among individuals in a way that did not change with age over the range of ages in these experiments. Second, the dual-task penalty was clearly associated with the position effect. The canonical coefficients showed that children who suffered a large penalty when the secondary recall task was imposed also experienced a smaller benefit when the correct response was from the last (easier) position. This finding lends support to the argument made above that the position effect in sentence completion may be related to the recency effect in free recall.

General Conclusions

Our experiments support the conclusion that sentence completion produces effects broadly parallel to those typically obeserved in sentence verification. Whatever cognitive processes produce the negation and comparator effects in verification were probably responsible for the similar effects we found in completion. However, despite the presence of similar effects, there may be an important difference concerning magnitude of effects. In verification, where the major variable is usually reaction time at nearly perfect accuracy, the extra time for negation is ordinarily much larger than the difference in time between comparators (Carpenter & Just, 1975; Trabasso, 1972). In the present experiments on completion, where the major variable was benchmark accuracy at intermediate levels between chance and errorless peformance, the effect of negation was distinctly and consistently smaller than the difference in accuracy between comparators. This reversal in the relative magnitude of the negation and comparator effects could be attributed to numerous procedural differences between the present experiments and published studies of verification. New studies that directly compare the two paradigms seem warranted.

An additonal conclusion is that individual differences in sentence completion are at least moderately stable. There is good reason to believe that such differences derive from variation in the efficiency with which different children executed elementary cognitive processes, including processes for managing memory loads and for performing an act of negation upon information held in a memorial representation. What remains in doubt is not whether there is stability in these differences but how much. It is possible that our correlations might have been greater if more extensive data involving many more trials had been obtained from each child.

References

Carpenter, P. A., & Just, M. A. Sentence comprehension: A psycholinguistic processing model of verification. Psychological Review, 1975, 82, 45-73.

Catlin, J., & Jones, N. K. Verifying affirmative and negative sentences. Psychological Review, 1976, 83, 497-501.

Clark, H. H., & Chase, W. G. On the process of comparing sentences against pictures. Cognitive Psychology, 1972, 3, 472-517.

Hunt, E. B. Mechanics of verbal ability. Psychological Review, 1978, 85, 109-130.

Keating, D. P., & Bobbitt, B. L. Individual and developmental differences in cognitive-processing components of mental ability. Child Development, 1978, 49, 155-167.

Kintsch, W. The representation of meaning in memory. Hillsdale, NJ: Erlbaum, 1974.

MacLeod, C. M., Hunt, E. B., & Mathews, N. N. Individual differences in the verification of sentence-picture relationships. Journal of Verbal Learning and Verbal Behavior, 1978, 17, 493-508.

Mosteller, F., & Tukey, J. W. Data analysis and regression. Reading, MA: Addison-Wesley, 1977.

Potts, G. R., Banks, W. P., Kosslyn, S. M., Moyer, R. S., Riley, C. A., & Smith, K. H. Encoding and retrieval in comparative judgments. In N. J. Castellan, Jr., & F. Restle (Eds.), Cognitive theory (Vol. 3). Hillsdale, NJ: Erlbaum, 1978.

Shoben, E. J. Choosing a model of sentence-picture comparisons: A reply to Catlin and Jones. Psychological Review, 1978, 85, 131-137.

Sternberg, R. J. Intelligence and nonentrenchment. Journal of Educational Psychology, 1981, 73, 1-16.

Trabasso, T. Mental operations in language comprehension. In J. B. Carroll & R. O. Freedle (Eds.), Language comprehension and the acquisition of knowledge. Washington, DC: Winston, 1972.

Trabasso, T. The role of memory as a system in making transitive inferences. In R. V. Kail, Jr., & J. W. Hagen (Eds.), Perspectives on the development of memory and cognition. Hillsdale, NJ: Erlbaum, 1977.

Wason, P. C. The processing of positive and negative information. Quarterly Journal of Experimental Psychology, 1959, 11, 92-107.

Table 1

Interaction concerning the Preposition Effect: Accuracy for After and Before with Letters and Words as Input (Experiment 1)

Preposition	Letters	Words	Sentences
after	.86	.62	.76
before	.43	.79	.63

Note. The tabled values are on a scale of proportion correct, computed by converting to this scale from adjusted benchmark accuracy on the logistic scale.

Table 2

Interaction concerning the Position Effect: Accuracy for Positions 1 and 3 as the Correct Position with Letters and Words as Input (Experiment 1)

Position	Letters	Words	Sentences
1	.80	.56	.70
3	.50	.84	.69

Note. The tabled values are on a scale of proportion correct, computed by converting to this scale from adjusted benchmark accuracy on the logistic sc scale.

Table 3
Interaction between Age and Adjective: Accuracy of Younger and Older Children for Same and Opposite Adjective (Experiment 2)

Adjective	Younger	Older	Mean
Same (faster)	.69	.75	.72
Opposite (slower)	.56	.59	.57

Note. The tabled values are on a scale of proportion correct, computed by converting to this scale from adjusted benchmark accuracy on the logistic scale.

Table 4
Interaction between Age and Position: Accuracy of Younger and Older Children for Positions 1 and 3 as the Correct Position (Experiment 2)

Position	Younger	Older	Mean
1	.59	.67	.63
3	.66	.68	.67

The tabled values are on a scale of proportion correct, computed by rting to this scale from adjusted benchmark accuracy on the logistic

Table 5
Canonical Correlations Within and Between Experiments

Variable	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Experiment 1							
Overall performance (1)	.38**	.32*	.38**	.19	.15	.09	.28*
Negation effect (2)		.21	.19	.26	.16	.16	.13
Experiment 2							
Overall performance (3)			.34*	.22	.30	.32*	.11
Negation effect (4)				.20	.30	.23	.10
Adjective effect (5)					.28	.24	.12*
Position effect (6)						.44**	.09
Dual-task penalty (7)							.04
Recall score (8)							

* $p < .05$
** $p < .01$