DOCUMENT RESUME

ED 108 180 CS 001 970

AUTHOR Harris, Wendy J.; Rohwer, William D., Jr.
TITLE Children's Memory for Between-Sentence and Within-Sentence Relations.
PUB DATE 75
NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30-April 3, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS *Child Language; Children; *Cognitive Processes; *Language Skills; Memorizing; *Memory; *Reading Research; Reading Skills; Recall (Psychological); Semantics; Sentences; Sentence Structure
IDENTIFIERS *Spatial Relationships (Sentences)

ABSTRACT

This study investigates children's semantic integration of sentence information as a function of instructions (form or substance), test sentence form (verbatim or paraphrased from acquisition story sentences), and story content (spatial or general relationships). After 144 fifth-grade children were presented with twelve short acquisition stories, they were asked to discriminate between true and false premise or inference recognition test sentences. Performance varied as a function of experimental condition, and discrimination for correct premises was much better than that for correct inferences. No consistent support was found for a view of constructive memory in children. (Author)

CHILDREN'S MEMORY FOR BETWEEN-SENTENCE AND WITHIN-SENTENCE RELATIONS

Wendy J. Harris and William D. Rohwer, Jr.

University of California, Berkeley

Research in language and memory has frequently addressed the issue of how sentences are understood and remembered. One approach to the issue, the constructive view (e.g., Barclay, 1973; Barclay & Reid, 1974; Bransford, Barclay, & Franks, 1972; Bransford & Franks, 1971; Johnson, Bransford, & Solomon, 1973), assumes that during storage subjects use the information in sentences to construct semantic descriptions of situations, objects, events, and the like. These descriptions consist not only of information extracted from the sentences, but also of the information produced as subjects assimilate the sentences to their prior knowledge. Thus, the product of these constructive storage operations is an indivisible compound of inferred information as well as information explicitly stated in the original sentences. If subjects habitually store information in this manner, subsequently they should be unable to distinguish between explicitly presented information (premises) and information constructed from it (inferences). Still, subjects should be able accurately to distinguish between information congruent and incongruent with what was presented, whether in premise or inference form.

These predictions were tested in a study by Bransford, Barclay and Franks (1972), who presented adult subjects with sets of three sentences describing spatial relationships among objects. In a forced choice recognition task, subjects were then asked to select from among a set of four alternative test sentences which they had actually heard. The four sentence choices were either valid or invalid premises or inferences. Subjects were expected to select valid inferences as frequently as valid

premises and to select both types of valid sentences more frequently than either of the invalid types. Consistent with a constructive view of sentence memory, these predictions were substantiated for test sentences having either a verbatim or a paraphrase relationship to the study sentences.

Two studies have extended the constructive model to children of age seven to eleven years. In a task similar to that used by Bransford, Barclay, and Franks (1972), Paris and Carter (1973) concluded that, like adults, second and fifth-grade children spontaneously integrate semantic information by constructing between-sentence relations during storage. In a true false recognition test, children of both age groups accepted valid inferences about as often as they accepted valid premise statements, and accepted both at a substantially higher rate than that for invalid sentences. This result was replicated and extended to pictorial stimuli by Paris and Mahoney (1974).

The finding that children appear unable to distinguish between valid inferences and explicitly presented premises is consistent with the contention that they, as well as adults, spontaneously construct between-sentence relations. However, before accepting this conclusion, several methodological issues raised by previous studies must be resolved. The present study was designed to address these issues.

First, in the two studies reported by Paris and his colleagues, subjects were instructed to remember the study sentences, and prior to presentation of the test sentences, they were told to accept a sentence if it were exactly the same as one of the sentences presented for acquisition. Such instructions emphasize surface linguistic forms of sentences. Thus, it may be important that in these and other studies (e.g., Bransford, Barclay, & Franks, 1972; Johnson, Bransford, & Solomon, 1973) the invalid test sentences all contained a word not presented in the original study sentences. It is conceivable

that subjects rejected invalid test sentences, not because of constructed semantic information, but because the sentences contained novel words. Since children were given instructions to pay attention to sentence form, rather than to meaning, there is the possibility that rejections were based upon memory for surface rather than semantic information. In support of this possibility, Paris and Mahoney (1974) found that when test sentences were paraphrases of the original study sentences, children were unable to discriminate between valid and invalid premise or inference sentences. Adults, on the other hand, were able to discriminate reliably among such paraphrase test sentences (Bransford, Barclay, & Franks, 1972). Thus, in the present experiment, test-sentence form (verbatim or paraphrase) was manipulated in conjunction with instructions (form or meaning) to determine which of the test-sentence features is functional in discrimination performance. Novel words were introduced in test sentences only in the paraphrase conditions, and they appeared consistently for all types of sentences in that condition.

A second potential problem in the Paris studies is that the four types of test sentences derived from a single acquisition story were presented contiguously in the recognition list. This same format was used in the Bransford, Barclay, and Franks experiment except that the method was forced choice instead of a recognition task. Because these sets of four sentences contained dependent information, it is possible that a subject's response to one sentence may have resulted from his response to the others or by information available in them. Thus, in the present experiment, the design of the recognition list was varied to see if performance is affected by such test sentence dependencies.

Finally, the results of the experiments with children are specific to relationships which were spatial in nature. To determine the generality of the experimental effects pertaining to constructive memory, sentence relationships in the present study were sampled from a wider variety of types.

To summarize, the present experiment was designed to determine the validity and generality of the claim that children naturally construct between-sentence and within-sentence relations. According to a constructive memory model, such subjects should be able to discriminate valid from invalid sentences, whether inferences or premises. To assess this prediction, memory for sentence information was investigated as a function of instructions, test sentence form, recognition list design, and type of sentence relationships.

METHOD

Subjects

A sample of 144 white, fifth-grade students were used in the study. Subjects were randomly selected from classrooms serving a high socioeconomic status residential area.

Materials

Twelve short stories were constructed for study materials. Half of these stories described spatial relationships among objects, and half of these stories described relationships of a more general nature (comparatives, for instance). An example of the latter type is:

Trains are heavier than cars. Cars remain upright in the wind.
Trees are taller than bushes. Bushes provide shelter from the rain.

Paraphrase versions of each of the stories were also constructed, e.g.,:

Trains weigh more than cars. Cars stay right side up in the wind.
Trees are higher than bushes. Bushes give protection from the rain.

Four types of test sentences were formed from each story, e.g.:

Valid premise - Trains are heavier than cars.
Invalid premise - Trains are taller than cars.
Valid inference - Trains remain upright in the wind.
Invalid inference - Trains provide shelter from the wind.

Several comments can be made concerning these materials. First, the test sentences formed from any story were in a verbatim relationship to one form of that story and a paraphrase relationship to the other form of that story; thus, test sentence form was varied. The two forms of the materials were counterbalanced, so that within each experimental group, half the subjects received one form for study while the other subjects received the other form.

Another feature of the materials is that since there were two concepts developed in each story, it was possible to falsify a statement by substituting a relationship from the opposite concept in the story (e.g., "taller" for "heavier"). So in the verbatim conditions, all of the words in the test sentences were drawn from the study materials. In addition, since the sentences were rendered invalid by substituting an opposite relationship, and since all test sentences were equally plausible in a pragmatic sense, subjects would have had to encode the original sentence meanings in order to make correct rejections.

Two recognition list designs were used. In each list, the four test sentence types were equally represented. The Within-stories design essentially replicated that used in the previous studies. In this design, the test list contained 32 sentences, and the four types of test sentences were selected from each of eight stories, so that there was dependent information on the list. The four stories not represented in a list were counterbalanced over subjects.

The Between-stories list design consisted of a 16-sentence test list, where at most two (non-dependent) premise statements were taken from a single story, and where all stories were represented in a single test list. One aspect of this design is that all possible sentences from all stories can be represented over a group of 12 subjects.

Procedure

Subjects were randomly assigned to experimental conditions and tested individually. A subject was told to listen very carefully to a group of short stories, because afterwards, he would be asked questions about what he had heard. A sample story was read, and the appropriate instructions were given for recognition sentences. Subjects in the Form instructions group were told to answer "yes" if the sentence they heard was exactly the same as one of the ones they had heard in the story, and "no" if it was not exactly the same. Subjects in the Meaning instructions group were told to answer "yes" if the test sentence had the same meaning as what they had heard in the story. A valid and an invalid premise test sentence were given for practice, presented in either verbatim or paraphrase relationship to the story, depending on the appropriate experimental condition. Responses to these practice sentences were then corrected, if necessary. When it was clear to the experimenter that the subject understood the directions, the 12 stories were presented via a portable Sony tape recorder at a normal speaking rate. Stories were identified by number on the tape. After a one-minute retention interval, during which the subject was reminded of his instructions, the list of test sentences was read. For each sentence, the subject orally responded yes or no, and then rated his confidence on a 3-point scale (very sure, kind of sure, not very sure). All answers were recorded by the experimenter.

Design

The entire experiment can be conceived as two separate studies following two different recognition list designs. To insure an equal number of observations for each type of recognition sentence, 48 Ss comprised the Within-stories design, and 96 Ss comprised the Between-stories design. Within each of these list designs, the two between-subjects factors were Instructions (form or meaning) and Sentence Form (verbatim or paraphrase). The two within-subjects factors were Information type (premises or inferences) and Relations (spatial or general).

RESULTS

Two dependent variables were used to assess discrimination of valid premises and inferences among the different experimental groups. One was the difference between proportion of valid sentences accepted and invalid sentences accepted (discrimination) and the second was the proportion of sentences accepted. Results are reported separately for each of these dependent variables.

Discrimination

The discrimination variable, a bias-free performance measure, was used to assess the effects of the factors of Instructions and Sentence Form, as well as to compare memory for premises and inferences within each of the experimental groups.

- - - - - - - - - - - - - -

Insert Table 1 about here

- - - - - - - - - - - - - -

Table 1 presents the results of premise and inference discrimination for spatial and general relations as a function of List Design, Instructions,

and Sentence Form. Discrimination of each type of test sentence (spatial and general premises and inferences) and discrimination of total premises and inferences for the two list designs were all analysed in a series of two-way analyses of variance, with Instructions and Sentence Form as the factors. Among all of these analyses, there were only two significant effects. One was a main effect for Instructions on general premises in the Within-stories design ($F = 5.48$, $df = 1/44$, $p < .05$), for which discrimination was better when subjects were given the Meaning instructions. The other effect occurred for Sentence Form on total inferences in the Between-stories design ($F = 4.44$, $df = 1/92$, $p < .05$). The means in Table 1 indicate that inference discrimination was much worse for paraphrase than for verbatim test sentences whereas premise discrimination was not affected by the difference in sentence form ($F < 1$). In general, no consistent effects of the Instructions and Sentence Form factors were found for the different types of test sentences, and discrimination of the test sentences varied considerably over the four experimental conditions.

The discrimination measures in Table 1 also provide a way to evaluate the predictions of a constructive model of sentence memory in children within and across the experimental conditions. According to such a model, discrimination of valid inferences should be about as high as discrimination of valid premises, since both types of sentences contain information semantically congruent with the acquisition story. Secondly, discrimination of both types of information should differ reliably from a chance level (zero).

Looking first at whether discrimination was equivalent for the premises and inferences, the mean scores show that with only two exceptions, premise discrimination was substantially higher than inference discrimination. One

of the exceptions was in the Meaning-Verbatim group in the Within-stories design, where discrimination between valid and invalid inferences was greater than discrimination between valid and invalid premises. This was consistent for both types of relations. The other exception was for spatial relations in the Form-Paraphrase group in the same design. Here, inference and premise discrimination were equivalent, but neither was significantly different from zero.

The means in Table 1 which differ significantly from a chance level ($p < .05$) are indicated with an asterisk. The results vary considerably between the two designs. For all items together in the Between-stories design, only in the Form-Verbatim group was inference as well as premise discrimination significantly different from zero, but the significant effect of inference discrimination was specific to the spatial items. In the other three groups, subjects were able reliably to discriminate valid premises, but not valid inferences.

In the Within-stories design, results from all items together indicate that in all but one group (the Meaning-Verbatim group), inference and premise discrimination were both significantly different from zero. Although the results of the Within-stories design indicated that premise discrimination was still generally better than inference discrimination, because of the result that in one of the groups this trend was reversed, and because of the likelihood that dependencies among sentences affected results in the Within-stories design, this design does not appear promising as a way to assess memory for within-sentence and between-sentence relations.

Finally, to determine how performance differed as a function of confidence assigned to a response, discrimination at the three levels of confidence was computed. This discrimination measure rarely exceeded chance at either

medium or low levels of confidence, and results of discrimination at high confidence generally paralleled results when confidence was ignored.

Proportion of Test Sentences Accepted

To make a direct comparison between the results of the present study and those obtained by Paris and his colleagues, performance was indexed in terms of the proportion of each type of test sentence accepted. Unlike the discrimination measure, this index allows a comparison of performance on valid and invalid test sentences. In Table 2, the proportions of premise

- - - - - - - - - - - - -

Insert Table 2 about here

- - - - - - - - - - - - -

and inference test sentences accepted are presented for the Form-Verbatim group as a function of List Design and Relations. The Form instructions and the Verbatim test sentences received by this group approximated the experimental conditions in the Paris and Carter (1973) study.

Comparing performance on the valid premise and inference test sentences, Paris and Carter found that subjects tended to accept almost as many inference statements as premises (difference between the two proportions was about .17), even though the inferences might have been rejected more frequently since they were never exactly the same as any sentences in the stories. The data in Table 2 indicate that the difference between valid premises and inferences accepted in the present study was consistently greater than the difference obtained by Paris and Carter, reaching a maximum for the general items in the Between-stories design (.61).

DISCUSSION

The results of this study provide little evidence that children spontaneously integrate semantic information across sentences. The fifth-

grade subjects were unable consistently to discriminate between valid and invalid inferences, as they could do with premises. Such a result is inconsistent with results reported by Paris and Carter (1973), and suggests that outcomes of previous studies are a function of a specific methodology, since the effects are not replicable over a variety of experimental conditions.

Another result which is inconsistent with a constructive view of sentence memory in children was that discrimination of inferences was significantly worse when the inference test sentences were paraphrased from the stories, while premise discrimination remained the same. This result is congruent with those obtained by Paris and Mahoney (1974), and viewed together, these two studies imply that children are not able consistently to construct between-sentence relations, and thus do not behave on this task in accord with the predictions of a constructive model of sentence memory. With reference to between-sentence relations, the only evidence in the present study that memory processes in children are characteristically constructive comes from the Within-stories groups and the Form-verbatim group in the Between-stories design. However, discrimination for inferences in the latter group was still below that for premises.

The successful discrimination of inferences in the Within-stories list design can conceivably be accounted for by the dependencies among the four types of test sentences all drawn from the same story. For one thing, on the test list, subjects were presented with sufficient information from the premise statements to correctly accept or reject the inference sentences, if the inference sentences appeared on the list after the premises. This could obviously have happened for only some of the stories, because sentences from stories were randomly ordered. For another thing, a valid inference and an invalid inference were presented from the same story, and a subject might thus realize that one of them should be accepted, and one

should be rejected. In general, both the inconsistent results and the interitem dependencies inherent in this design suggest that it has little, if any, future utility.

Another methodological recommendation for future research concern the manner in which recognition sentences are rendered invalid. In previous studies, the invalid premise test sentences were rarely accepted as having been seen before (the proportion of false alarms by fifth-grade subjects in the Paris and Carter study was .10), whi[illegible] present study, the false alarm rate was substantially higher (about .45 for both list designs). Paris and Carter assumed that subjects rejected invalid sentences for semantic reasons, that is, because the sentence information was incongruent with what they had heard before. However, such rejections may have been based upon lexical cues, since sentences were rendered invalid by replacing original prepositions with ones the subjects had not seen before. Such novel words could have been used as cues to reject the test sentences, especially since Form instructions were given. By contrast, in the present study, none of the test sentences in the Verbatim conditions contained novel words, since invalid sentences were formed by substituting a different, but familiar relationship, and the false alarm rate was substantially higher. The major implication is that to assess the information subjects have encoded semantically, the test materials should not permit discrimination based solely on lexical cues. Katz (1974) has offered a similar argument for extending the results of Bransford and Franks' (1971) experiment to a Meaning instructions condition.

The finding that discrimination of valid and invalid spatial inferences

was better than discrimination of the general inferences calls for explanation. One hypothesis is that subjects were using an imagery strategy appropriate for the spatial but not the general items. For at least two reasons, however, this hypothesis seems implausible. First, children of this age do not customarily use such a strategy without having been instructed to do so (e.g., Rohwer, 1973). Second, subjective reports from several of the subjects indicated that few of them tried to imagine the relative locations of the items to be learned, and those who did so did not do it consistently. A more likely interpretation has to do with the probability that an inference containing spatial relationships can actually be formed. That is, if a subject is told that A is to the left of B, and that B is above C, making the inference that A is also above C follows logically from the premises. The same cannot be said of the more general relationship items where in order to make the appropriate inference, a subject must have available certain implicit assumptions about the state of the world. Thus, as children get older, and their knowledge about the world increases, they should be more likely to make such inferences. An appropriate developmental design could be used to assess this interpretation.

References

Barclay, J. R. The role of comprehension in remembering sentences. Cognitive Psychology, 1973, 4, 229-254.

Barclay, J. R., & Reid, M. Characteristics of memory representations of sentence sets describing linear arrays. Journal of Verbal Learning and Verbal Behavior. 1974, 13, 133-137.

Bransford, J. D., & Franks, J. J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.

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

Johnson, M. J., Bransford, J. D., & Solomon, S. K. Memory for tacit implications of sentences. Journal of Experimental Psychology, 1973, 98, 203-205.

Katz, S. Role of instructions in abstraction of linguistic ideas. Journal of Experimental Psychology, 1973, 98, 79-84.

Paris, S. G., & Carter, A. Y. Semantic and constructive aspects of sentence memory in children. Developmental Psychology, 1973, 9, 109-113.

Paris, S. G., & Mahoney, G. T. Cognitive integration in children's memory for sentences and pictures. Child Development, 1974, 45, 663-642.

Rohwer, W. D., Jr. Elaboration and learning in childhood and adolescence. In H. W. Reese (Ed.), Advances in child development and behavior. New York: Academic Press, 1973.

Footnotes

[1] This research was supported, in part, by a grant from the National Institutes of Health (HD03869). The report was prepared at the Institute of Human Learning, which is supported by grants from the National Institutes of Health.

[2] We wish to acknowledge the cooperation of the students, faculties, and principals of schools in the Orinda School District.

Table 1

Discrimination as a Function of List Designs, Instructions, and Sentence Form

Instructions	Sentence Form	Information Type	Within-Stories Design			Between-Stories Design		
			Spatial	General	Total	Spatial	General	Total
Form	Verbatim	Premise	.21*	.54*	.37*	.29*	.48*	.39*
		Inference	.06	.21*	.14*	.21*	.08	.15*
	Paraphrase	Premise	.08	.36*	.22*	.14	.36*	.25*
		Inference	.08	.20*	.14*	-.04	-.06	-.05
Meaning	Verbatim	Premise	.02	.19*	.10	.48*	.30*	.39*
	Paraphrase	Inference	.21*	.25*	.23*	.08	.12	.11
	Paraphrase	Premise	.19	.27*	.23*	.48*	.31*	.39*
		Inference	.10	.23*	.17*	.04	.04	.04

* Mean difference in proportions significantly different from chance, or zero (α = .05)

Table 2

Proportion of Test Sentences Accepted for the Form-Verbatim Group as a Function of List Design, and Relations

Information Type	Within-Stories Design			Between-Stories Design		
	Spatial	General	Total	Spatial	General	Total
Valid Premise	.67	.92	.79	.73	.92	.82
Invalid Premise	.46	.38	.42	.44	.44	.44
Valid Inference	.42	.50	.46	.42	.31	.36
Invalid Inference	.35	.29	.32	.21	.23	.22