DOCUMENT RESUME

ED 028 902 RE 001 704

By-Bormuth, John R.; And Others

Children's Comprehension of Between and Within Sentence Syntactic Structures.

Pub Date Feb 69

Note-15p.; Paper presented at American Educational Research Association conference, Los Angeles, Feb. 5-8, 1969

EDRS Price MF-$0.25 HC-$0.85

Descriptors-*Grade 4, Questioning Techniques, *Reading Comprehension, *Sentence Structure, *Syntax

Three broad categories of comprehension skills (sentence, anaphora, and intersentence syntax) were divided into a total of 55 separate skills. Two different sentences or sentence pairs were written to incorporate each of the structures studied, and a four- or five-sentence paragraph was then written to incorporate each of these. Four question types (rote, transform, semantic substitute, and compound) were used to test comprehension of the structures; therefore, four test forms using these different question types were made for each paragraph. The subjects, fourth graders from an inner city, a suburban, and a rural school, were randomly divided so that 60 students responded to each question. The mean percentage of the students correctly answering the sentence comprehension questions was 73 percent; the anaphora comprehension questions, 77 percent; and the intersentence questions, 58 percent. The structures identified seemed to represent homogenous classes of behavior since the variation between questions measuring different skills was significantly greater than the variation between items measuring the same skill. The fact that the structures and question types differed significantly in difficulty was also taken as evidence that many of these skills may be hierarchically related. Tables and references are included. (CM)

AERA

Children's Comprehension of Between and Within Sentence Syntactic Structures

John R. Bormuth, University of Chicago
John Manning, University of Minnesota
Julian Carr, University of Minnesota
David Pearson, University of Minnesota

Since much of the knowledge contained in the school's curriculum is transmitted through the medium of written language, the failure or success of the educational enterprise is heavily dependent upon how well students are able to comprehend the language in their instructional materials. Unfortunately, the testing procedures in current use are unable to provide the information necessary for determining how well students are able to understand the syntactic structures by which language signals information. Consequently, we have little knowledge upon which to base the design of instruction for teaching these language comprehension skills.

This study represents a preliminary attempt to identify these skills and then to determine first if the skills identified represent homogenous classes of behaviors, second if the skills might be hierarchically related, and third what is the general level of performance on these skills displayed by children in grade four where comprehension instruction is normally begun in earnest.

This study contrasts in three important respects with other attempts to measure students' comprehension of language. First, it addresses itself to the problem of constructing an instructional theory of comprehension rather than to the construction of a theory which describes the processes involved in the comprehension of syntactic structures. A psychological account of the comprehension processes is, of course, relevant to the instructional theory of comprehension, but it falls short of being sufficient for the design of instruction. The contrast between the two types of theory can be understood by considering the diagram in Figure 1 showing the operational unit of instruction. A psychological theory of language comprehension attempts only to give an account of the behaviors represented in block C. The experiments used to construct this theory utilize test tasks as in block B and responses as in block F. But the form of the test tasks are selected and systematically varied in a way that permits the experimenter to infer just the nature of the behavior represented in block C.

A psychological theory is not sufficient for instructional theory, for it does not account for all the important sources of variation in the actual operational setting of instruction. An inescapable fact of instruction is the necessity of using test-like tasks. This forces the instructional theorist to take fully into account variation attributable to the behaviors represented by blocks D and E, as well as the behaviors represented by block C. The test-like tasks are essential in instruction because they provide (a) a means of practicing the student in the behavior, (b) a source of feedback to the student which informs him of the correctness of his behavior, and (c) a source of feedback to the instructor giving him the information necessary for managing and altering the instruction. Without these test-like tasks, there is no way to ascertain what, if anything, is being learned or even if the student is doing anything at all relevant to the instructor's intentions. Hence, the instructional theorist must concern himself also with the effects introduced by the test-like tasks actually used in instruction.

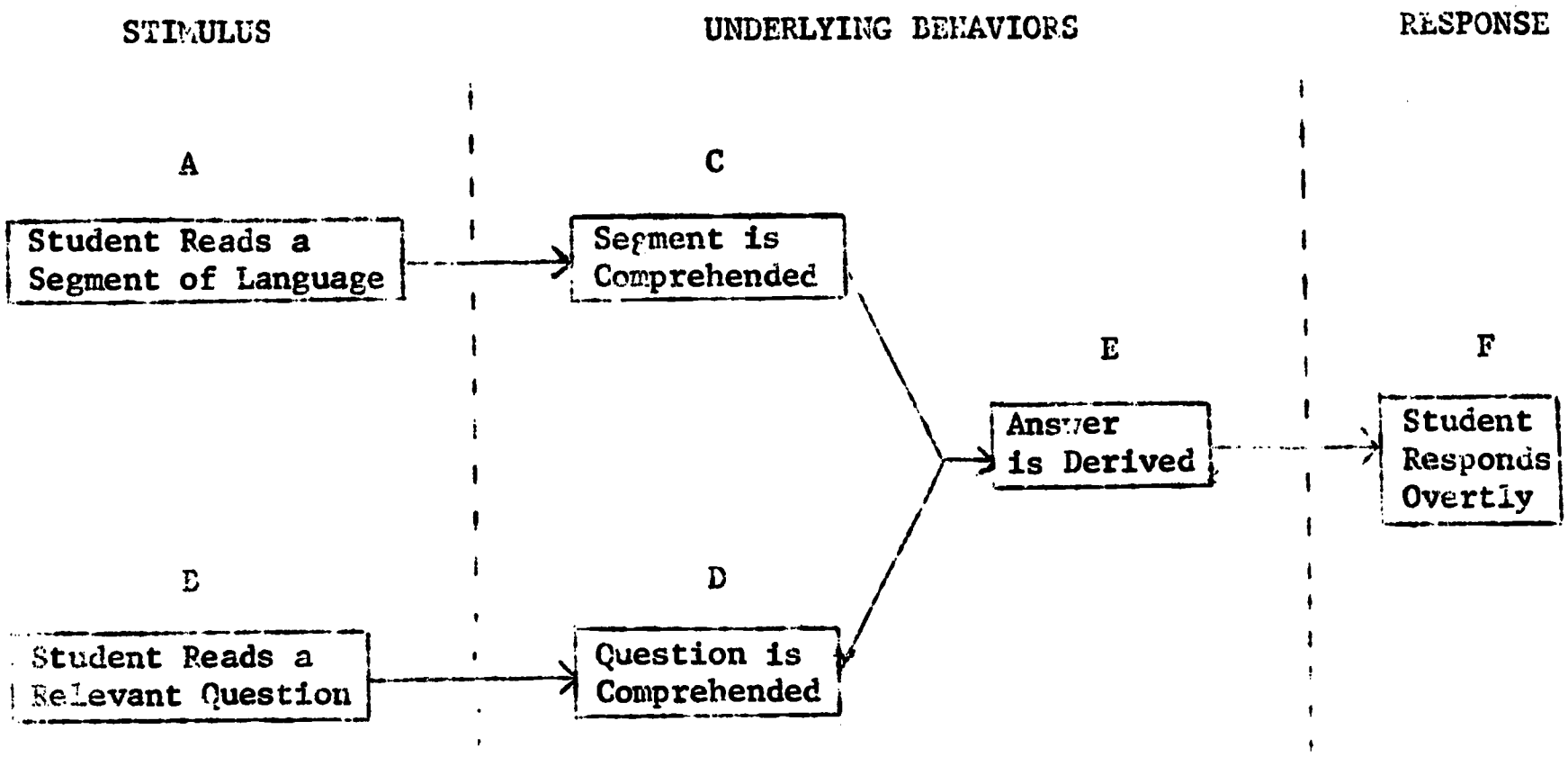

Figure 1. The operational unit of instruction

It should also be understood that an instructional theory is not necessarily a by-product of the construction of a psychological theory. The psychological theorist is free to select a broad range of test-like tasks regardless of whether they necessarily force the subject to emit the desired underlying behavior. For example, psychologists studying sentence processing frequently employ tasks which require the subjects to memorize and repeat sentences. While these tasks may be quite adequate to reveal the phenomenon the psychological theorist is interested in, he would certainly not claim that children could be taught to comprehend sentences merely by memorizing them. Consequently, the psychological theorist leaves unexamined the effects of many of the test-like tasks which the instructional theorist must examine if he is to select and sequence the content of language instruction.

This discussion should also make it clear that an instructional theory of comprehension is incomplete unless it includes not only an account of the difficulty of, say, a syntactic structure but also an account of the difficulty of the different test-like tasks used to teach that structure. In actual instructional practice, a student can be said to have learned a behavior only if he can exhibit a correct response to all of the test-like tasks ordinarily used to test that behavior. Consequently, the task of the instructional theorist is to provide the information by which instruction on the test-like tasks, also, can be designed.

The second major contrast between this and other studies of language comprehension is that it rejects the traditional practices followed in taxonomizing cognitive skills. The definitions in those taxonomies depend primarily upon the introspections of a test writer rather than employing the standard scientific practice of referring to publicly observable operations. For example, test writers (Bloom, 1956, for example) often refer to items which test comprehension of important facts, ability to evaluate, and the like. Those authors do not provide definitions of these items by which another investigator or a teacher could identify, say, a fact much less operations by which others could discriminate important from less important facts.

The method used to define the items in this study is sufficiently operational that these identical items and items of the same types can be derived by any person with some competence in linguistics or even, potentially, by a properly programmed computer. Thus, this study, as should all studies in instructional theory, can claim that its results are replicable across all items and structures of the same types as those studied here, without the need to consider the possibility that the introspections which the next investigator brings to the item writing task might not match the introspections of the original investigators.

The third contrast is that this study uses a finer grained analysis of what is often termed the literal comprehension skills than has heretofore been used. In the most analytic studies to date, the practice has been to provide measures of at most a half dozen skills. The present analysis included skills in the three broad categories of sentence, anaphora, and intersentence syntax comprehension and then further broke these down into a total of 55 skills.

Procedure

Taxonomy of Comprehension Skills: In this analysis a comprehension skill is defined as the ability to respond correctly to a wh- question which deletes one of the immediate constituents of a syntactic structure. Consider, for example, the sentence The small boy rode the black horse. which contains noun phrase structures of the form adjective-plus-noun, small boy and black horse. For all sentences containing such structures a wh- question can be derived roughly by deleting the adjective, replacing it with the appropriate wh- pro word, and then, if it is not already there, shifting the pro word to the front of the sentence. These operations obtain questions like Which boy rode the black horse? and Which horse did the small boy ride? Similar questions can be derived from each of the different types of structures appearing in English sentences. The label of the skill presumably tested by these questions is derived from the structure the question tests. These particular questions, for example, are said to test the comprehension of prenominal adjectives.

The original taxonomy of structures appearing in English sentences identified 52 structures. (For a detailed listing and definitions of all structures referred to in this study see Bormuth, 1969, and Menzel, 1969.) From this set, the 25 structures which, upon inspection, seemed least likely to be understood by all fourth grade students were selected. (Table 2 gives the label and an example of each of these structures.)

The taxonomy of anaphora used in the study lists 14 different types. An anaphoric expression is a pronoun-like structure which shortens or substitutes for an expression which is usually antecedent to it and which has the same referent as the antecedent. Pronouns are one type of anaphoric expression, but so is the expression This boy in The small boy rode the horse. This boy is a good rider. which shortens the antecedent, the small boy. A segment of any size may serve as the antecedent of an anaphora, from a single word to a whole chapter, but this study included only those simple anaphora having an antecedent fully contained in the sentence immediately preceding the one containing the anaphoric expression. (Table 3 shows the labels and examples of these anaphora.)

The taxonomy of intersentence syntactic structures consisted of the 16 categories shown in Table 4. The relative sequential positions of sentences and conjunctive expressions signal information about the relationships of the sentences to each other. For example, in the appropriate context the order of the sentences The boy fell off the horse. He broke his arm. signals that the boy's fall from the horse caused the breaking of his arm. But when the order of these sentences is reversed, the causality is also reversed permitting the paraphrase The boy's breaking of his arm caused his fall from the horse.

Four question types were used to test comprehension of most of the structures studied. Since detailed definitions are given elsewhere for each of these question types (Bormuth, 1969), only general descriptions will be given here. A rote question is derived by deleting one of the members of the structure to be tested, replacing it with the appropriate wh- pro word and then shifting the wh- pro word to the front of the sentence. Deriving these questions also often requires the syntactic transformations known as do insertions, flip-flops (Thomas, 1965), and the like. These details are ignored here for the sake of brevity. Table 1 shows examples of rote questions.

Transform questions are derived by first deriving a mediating sentence and then deriving a rote question from the mediating sentence. The mediating sentence is actually a paraphrase of the original sentence obtained by performing either a cleft or passive transformation on the original sentence. For example, either of the sentences It was the boy who rode the steed. or The steed was ridden by the boy. could have served as mediators for the transform questions testing sentence comprehension in Table 1.

A semantic substitute question is also derived through the use of a mediator sentence. To derive a mediator sentence, one or more synonymous terms are substituted for terms in the original sentences. The term substituted may be symmetrically related to the term in the original sentence in the sense that the two terms may be mutually substituted for each other without altering the referential meaning of the context. The pair break and fracture seem to be symmetrically related in the context broken arm and fractured arm. The substituted term may also refer to a more inclusive concept which hierarchically dominates the concept referred to by the term in the original sentence. The pair horse and steed represent such a pair since all steeds are horses but not all horses are steeds.

Table 1

Illustrations of Question Types

SENTENCE COMPREHENSION QUESTIONS	
Original Sentence:	The boy rode the steed.
Rote:	Who rode the steed?
Transform:	By whom was the steed ridden?
Semantic Substitute:	Who rode the horse?
Compound:	By whom was the horse ridden?
ANAPHORA COMPREHENSION QUESTIONS	
Original Sentences:	The boy fell off the steed. He fractured his arm.
Rote:	Who fractured his arm?
Transform:	Who was it who fractured his arm?
Semantic Substitute:	Who broke his arm?
Compound:	Who was it who broke his arm?
INTERSENTENCE SYNTAX QUESTIONS	
Original Sentences:	The boy fell off the steed. He fractured his arm.
Rote:	What caused the fracturing of the boy's arm?
Reversal:	What did the boy's fall from the steed cause?
Semantic Substitute:	What caused the breaking of the boy's arm?
Compound:	What was the breaking of the boy's arm caused by?

Hierarchically related terms are not symmetrically substitutable. When a hierarchically dominant term is substituted to form the mediator sentence, useful questions can be derived. But when hierarchically subordinate terms are substituted to form the mediator sentence, the answers to the questions derived are indeterminate. For example, given the sentence The boy fell off the horse. and the derived question Who fell off the steed? it is impossible to answer the question because it is uncertain whether the sentence referred to steeds, nags, or some other subset of the concept horse.

Compound questions represent questions derived by applying both the semantic substitute and the transform operations to sentences.

Questions testing comprehension of anaphoric and intersentence syntactic structures require the construction of mediator sentences by the use of an embedding step for their derivation. The anaphora is questioned by embedding the antecedent into the sentence containing the anaphoric expression and then deriving a wh- question which deletes the portion of the antecedent not appearing in the anaphoric expression. For example, from the pair of sentences The very small boy rode the horse. This boy was young. we obtain the mediator sentence The very small boy who rode the horse was young. in which we may replace either very small, who rode the horse or both with a wh- pro word. The mediator sentences used to derive intersentence syntax questions are formed by nominalizing the pair of sentences involved, inserting the appropriate conjunctive verb between them, and then replacing one of the nominalized sentences with the appropriate pro word. The questions in Table 1, for example, were derived from the mediator sentence The boy's falling off the horse caused the fracturing of the boy's arm. Note that instead of using transform questions to test the intersentence structures a reversal question was formed which deleted the sentence not deleted by the rote question.

Intersentence syntax, transform, and semantic substitute questions permit options in deriving the mediator sentences. In this study the nominalization and sentence paraphrasing transformation options were randomly selected for each transform and intersentence syntax question. The semantic substitutes were selected by selecting a symmetrically related term if one existed, and if it did not by selecting the hierarchically related term which both met the vocabulary constraints of the design and seemed to immediately dominate the original term in its hierarchy.

This manner of deriving questions produces questions of the completion type which have traditionally introduced problems of replicable scoring. A theory of scoring was developed (Bormuth, 1969) and then applied to the evaluation of the responses observed in this study. Since the theory and its results will be described in detail elsewhere (Carr, Pearson, and Boesen, 1969) only its gross features will be given here. A response was scored correct if it was the exact phrase replaced by the wh- pro word in the sentence or mediator sentence from which the question was derived. It was also scored correct if it was a correct semantic substitute or a grammatically correct anaphora of the phrase deleted. When this scoring theory was applied to the 26,400 responses in this study, all but roughly .002 of the responses were classifiable as clearly either correct or incorrect. That is, nearly all discrepancies between scorers were resolved as a scorer's error in "computing" the response rather than as insoluble differences in judgments.

Materials: Two different sentences or sentence pairs were written to incorporate each of the structures studied, and a four or five sentence paragraph was then written to incorporate each of these. In order to reduce syntactic complexity to a minimum, a rule was followed which permitted no sentence to contain more than

one embedded structure. To reduce the effects of vocabulary difficulty, all words were selected from a list of words which had been found to be easy to children (Dale and Chall, 1948) and then each paragraph was further revised if it seemed to contain an unusual usage of those words.

Each paragraph was printed with its question appearing immediately below it. Below the question three blank lines were provided for the student to write his answer on. Since there were four question types for each paragraph, four test forms were made for each paragraph, each containing the same paragraph but a different ~~structure~~ question type. A test booklet was compiled by randomly selecting one of the four test forms for each of the 110 paragraphs. The order of presenting the structures was rotated to counterbalance order effects. Thus, every subject was tested on every paragraph but no subject responded to more than one question type for a given paragraph.

Test Administration: The subjects were drawn in roughly equal numbers from the fourth grades of three schools: an inner city, integrated school, a suburban school and a rural school. All the schools were in the upper midwest area. The test booklets were randomly assigned to students and the tests administered in the regular classroom setting. Although the tests were administered to somewhat more students, the groups were randomly reduced so that there remained exactly 60 students responding to each question. The students were permitted all the time they needed to complete the tests.

Results

The percentages of students responding correctly to each sentence structure are shown in Table 2. The structures are ranked from easiest to most difficult using the percentages averaged across all four question types. These percentages were analyzed using a two factor, four question type by 25 structures, analysis of variance design in which the two examples of a structure provided two replicates in each cell. Both, the between structures variance (F, with 24 and 100 d.f., = 7.19) and the between question types variance (F, with 3 and 100 d.f., = 5.63) were significant at the .01 level and the interaction (F = .75) was not. Rote questions were easiest with a mean of .77 followed by transform questions having a mean of .71, semantic substitute questions having a mean of .69, and compound questions having a mean of .67.

Table 3 shows the anaphora structures similarly ranked. The between structures variation (F with 13 and 56 d.f. = 3.18) was also significant at the .01 level but the between question type variation did not reach significance (F with 3 and 56 d.f. = 1.03). The interaction was, again, less than unity (F = .46).

The same pattern of results was obtained from the analysis of the intersentence category of syntactic structures. The variation between structures (F with 15 and 64 d.f. = 4.82) was significant at the .01 level but neither the variation between question types (F with 3 and 64 d.f. = 1.59) nor the interaction (F = .95) was significant. Table 4 shows the difficulties of these structures.

Table 2

Percentage of Students Responding Correctly to Questions Testing Each of the Sentence Structures

Structure	Percent Correct
Comparative, unequal (Joe runs faster than Bill.)	88.3
Nominal Compound, noun + verb + preposition + noun (Hammer blow, a blow struck with a hammer)	87.9
Nominalization, possessive + verb + ing (His going came as a surprise.)	87.0
Prenominal Adjective (The tall boy is Joe.)	85.5
Subordinate Sentence, causal (We came because we smelled lunch being served.)	83.3
Relative Clause, with deletion (The man working in the yard is the owner.)	80.9
Nominalization, verb + ing (Finding him was easy.)	80.6
Adjectival Prepositional Phrase (The man with the hat manufactures cans.)	80.0
Subordinate Sentence, After (After we entered, the play began.)	79.8
Nominalization, factive (The fact that he came surprised us.)	78.0
Subordinate Sentence, although (Although it rained recently, the ground remains parched.)	77.3
Nominal Compounds, verb + ing (Washing machine, The machine is for washing clothes.)	77.3
Subordinate Sentence, purposive (In order to make certain, we asked a second time.)	76.3
Relative Clause, appositive (Mr. Joseph, who is our mailman, retired.)	73.7

Table 2 (continued)

Structure	Percent Correct
Nominal Compound, someone operates (Elevator operator, Someone operates an elevator.)	71.7
Subordinate Sentence, conditional (If we don't hurry, we'll miss the show.)	70.5
Nominal Compound, noun + preposition + noun (Ranger station, a station for rangers)	69.3
Nominalization, for-to (For us to find him was difficult.)	69.6
Nominal Compound, preposition + noun (Potato dumpling, The dumpling is made from potatoes.)	67.4
Relative Clause, without deletion (The man who has been working in the yard is the owner.)	67.0
Adjective Compliment (He is clever to go.)	66.2
Subordinate Sentence, before (Before we arrived, people had already been seated.)	65.1
Subordinate Sentence, tense shift if clause (If you had some money, you would buy some.)	61.4
Subordinate Sentence, simultaneous (As we entered, the curtain rose.)	56.1
Comparative, equal (Joe runs as fast as Bill.)	28.1

There were also differences among the categories of structures. The mean percentage of the students correctly answering the sentence comprehension questions was 73 per cent; the anaphora comprehension questims was 77 per cent; and the intersentence questions was 58 per cent. The intersentence category of questions was significantly more difficult than both the sentence questions (t with 40 d.f. = 3.83) and anaphoric questions (t with 30 d.f. = 4.18). The

Table 3

Percentage of Students Responding Correctly to Questions Testing Each of the Anaphoric Structures

Structures	Percent Correct
Pro-clause, so (Joe may go. If so, we will ....)	86.6
Pro-adverb (He works in the cellar. It is cool there.)	83.2
Relative Pronouns (The man who lives next door makes ....)	82.8
Pro-verb, so - do (Joe likes tennis. So does Bill.)	82.8
Deleted Modifier (The small boy came. This boy is ....)	82.4
Noun Phrase Demonstrative (The black horse belongs to Joe. That is his ....)	81.5
Numeric Pronominal (Several men went fishing. Two caught ....)	81.0
Inclusive Pronoun (Joe, Bill, and Mary went to the show. All enjoyed ....)	80.5
Deleted Noun (There are ripe and green apples. The green (...) are mine.)	78.6
Pro-verb, so - be/have (Joe is sick. So is Bill.)	76.1
Negated Pronoun (Bill and Joe went shopping. No one bought ....)	67.4
Clause Demonstrative (Joe is stuck in the mud. This leaves us ....)	66.3
Semantic Substitute (Those steel towers are antennas. These objects are ....)	65.5
Personal Pronouns (Joe left the room. He had ....)	64.5

Table 4

Percentage of Students Responding Correctly to Questions Testing Each of the Intersentence Structures

Structures	Percent Correct
Effect (The gear slipped off. The machine stopped.)	79.0
Cause (The machine stopped. The gear had slipped off.)	71.8
List (Three came. Joe, Bill, and Mary were the ones.)	68.9
And (Joe is a good hunter. He also skis well.)	67.3
Label (Joe skis well. He is Bill's brother.)	66.1
But (Bill said there was too much. Joe said there wasn't.)	65.9
Explanatory (Joe quit the team. He did not get to play enough.)	64.4
However (The gear slipped off. Surprisingly, the machine continued to run.)	63.5
Or, exclusive (Joe may have taken it. But it could have been Bill.)	56.6
Parenthetical (Joe had a cold. Colds seem common, don't they? His mother kept him home.)	52.9
Example (Mammals are warm blooded. Men and dogs are mammals.)	50.3
Although (The machine continued to run. This was inspite of the fact that the gear ....)	49.9

Table 4 (continued)

Structures	Percent Correct
Before (They had been working. Now they just stood there.)	48.3
Rule (Men and dogs are mammals. Mammals are warm blooded.)	46.9
While (Joe held the paper in his left hand. With his right hand he ....)	44.0
After (Joe found the others. He had looked everywhere for them.)	33.0

difference between sentence and anaphoric questions exhibited a t of less than unity. However, it should be recalled that, while the intersentence and anaphora structures tested exhausted their respective taxonomies, only those sentence structures were tested which seemed subjectively to be the more difficult ones in that category.

Discussion

By far the most startling result of this study was the fact that large proportions of the students were unable to demonstrate a comprehension of the most basic syntactic structures by which information is signaled in language. The success or failure of the educational process depends heavily upon the students having mastered these language comprehension skills. Yet, if these data are to be believed, large portions of the students tested were unable to demonstrate anything approaching such a mastery. And efforts to extrapolate these results to the general population lead only to a still grimmer picture since the school personnel claimed that the groups tested in this study ranged from average to well above average when their scores were compared to the national norms on the reading comprehension tests used by the schools. Consequently, these results deserve careful criticism.

The most common fallacy educators fall into in rationalizing test results grows from the fact that a conceptual distinction can be made between a student's having acquired a basic capability and his having acquired the instrumental capabilities

required to exhibit the basic capability on a test-like task. If this distinction can be accepted as meaningful, then it becomes a trivial matter to explain away results such as those reported here simply by arguing that the students quite possibly had acquired the basic competencies but that the data merely reflected the varying degrees to which the students had mastered the instrumental competencies necessary to demonstrate their basic competencies when faced with a test-like task.

While this distinction between basic and instrumental competencies is quite useful for some purposes, its use in this instance is fallacious because it refers to a statement which is meaningless in the most fundamental scientific and practical sense. A publicly verifiable claim, the only kind of claim having either scientific or practical significance, cannot be made that a student has gained a basic competency until that increment is observable in his publicly observable or overt behavior, and the situation which calls forth that behavior can always be conceptually regarded as involving instrumental competencies unique to it. Consequently, when a practitioner recognizes this conceptual distinction he is forced to forego the use of all practice exercises which provide practice for the student and feed-back to himself and the student. In actual practice, the unit of content must include not only the basic competency but also all of the instrumental competencies involved in the testing of the basic competency. Restated in the vernacular of the classroom, teachers must not only teach for a test, they must teach for every type of test likely to be used to elicit a demonstration of that competency.

So the interpretation of the present results is not affected by the argument that instrumental skills were involved, since no operational distinction of this sort is meaningful. However, it is reasonable to ask if the instrumental skills called for by the test-like tasks used in this study were so unreasonably difficult that they masked the basic competencies of the students. This seems unlikely to have occurred. The wh- questions used in this study are among the most common devices for testing competency. And careful attention was paid to holding the vocabulary and syntactic complexity of the materials to a minimum. Furthermore, no time limits were imposed either explicitly or implicitly on the testing situation and the tests were even arranged to conveniently facilitate the child referring back to the paragraph as he attempted to answer each question.

Thus, these data may actually over estimate the performance of students in actual instructional situations rather than underestimating it. For example, in a recent study (Bormuth, 1968) evidence was presented which indicated that only the most able elementary school children can gain information from the verbal instructional materials they are required to study and that even these students seem able to gain information only from the easiest of their materials.

Turning to the second problem investigated, it seems that the comprehension skills defined in this analysis represent fairly homogenous classes of behavior. This was shown by the fact that there were significant differences both between the difficulties of the major categories of structures and between the structures within the major categories. It is, of course, impossible from these data to certify that any two of the comprehension skills studied are indeed different behaviors. Such an analysis will require the use of many more test items of each

kind with careful attention being given to the effects the placement of a structure has on its difficulty. In this study these effects were randomized in so far as these placement options could be identified.

The difficulty ranking of skills is of fundamental importance in the design of instruction, since those rankings often reflect hierarchical relationships among the skills and thereby reveal the order in which skills should be taught. The fact that there were significant differences both among question types and among the structures within each major category suggests that such hierarchies exist. Interestingly enough, one of these appeared to run counter to the predictions which would have been made from transformational grammar. The transformational theory of grammar asserts that prenominal adjectives are derived from relative clauses with deletions which, in turn are derived from relative clauses without deletions. This would make prenominal adjectives the most complex and, therefore, the most difficult structure to comprehend. As it turned out, this order was exactly reversed. However, this result should be regarded as very tentative since this study was not designed to discriminate sharply between individual pairs of structures.

There were also indications that the major categories of structures may be hierarchically related as shown by the fact that they differed in difficulty. This ordering of difficulty was roughly the same as one would derive from linguistic theory. The expected ordering was that sentence structures would be easiest to comprehend, anaphora would be second, and intersentence structures the most difficult. The apparent reversal of this order occurring between sentence and anaphoric structures is likely to have been due to the fact that most of what subjectively seemed to be the easiest sentence structures were excluded from the study in the apparently mistaken belief that virtually all of the students would have demonstrated perfect performance on the items testing them.

There were also differences among the question types. While those differences were large enough to be considered significant only in the case of the sentence structures, all of the differences were in the expected directions. That is, rote questions were the easiest and compound questions the most difficult, while transform and semantic substitute questions ranged somewhere between the rote and compound questions. This outcome was expected because the procedures for deriving the questions themselves involve a hierarchy of operations.

Summary and Conclusions

This study represents a preliminary attempt to identify and operationalize the skills employed in comprehending the information signaled by the syntactic structures of language and then to determine if the skills identified represent homogenous classes of behavior, whether these skills might be hierarchically related, and finally what is the general level of performance on these skills displayed by fourth grade children.

From a taxonomy of 52 types of sentence structure, the 25 judged to be the most difficult were selected for testing. Also all 16 of the intersentence and all 14 of the anaphoric structures identified were tested. To do so two sentences containing a structure were written and each embedded in its own paragraph. Then four types of questions were constructed to test the structure being tested. Each structure with its questions was then used to test 60 fourth grade children.

The most startling result was the fact that large proportions of the children were unable to demonstrate a comprehension of even these basic structures by which information is signaled indicating that this deficiency may constitute a serious impediment to the efficiency of instruction. The structures identified seemed to represent homogenous classes of behavior since the variation between questions measuring different skills was significantly greater than the variation between items measuring the same skill. The fact that the structures and question types differed significantly in difficulty was also taken as evidence that many of these skills may be hierarchically related.

References

Bloom, B. S. (Ed.), Taxonomy of Educational Objectives, Handbook I: Cognitive Domain, New York: David McKay, 1956.

Bormuth, J. R., On the Theory of Achievement Test Items, Chicago: University of Chicago Press (in press).

Bormuth, J. R., "The Effectiveness of Current Procedures for Teaching Reading Comprehension," paper delivered at Convention of National Council of Teachers of English, Milwaukee, November 30, 1968.

Bormuth, J. R., "An Operational Definition of Comprehension Instruction," paper read at Symposium on Reading as a Psycholinguistic Process, February, 1967.

Carr, J. M.; Pearson, D.; and Boesen, M. "Reliability of a Procedure for Scoring Completion Test Items," (in progress).

Dale, E., and Chall, Jeanne S., "A Formula for Predicting Readability," Educational Research Bulletin, 27: 11-20, 1948.

Menzel, P., "Linguistic Bases of the Theory of Writing Items for Instruction Stated in Natural Language," in J. R. Bormuth, On the Theory of Achievement Test Items, Chicago: University of Chicago Press (in press).

Thomas, O. Transformational Grammar for the Teacher of English, New York: Holt, Rinehart, and Winston, 1965.