To examine the effect on comprehension of mistakes made while reading, a three-experiment study was conducted. A secondary aim was to determine whether the unskilled reader would make hypotheses about what words in the text should be. In the first test, 80 nine-year-old students with average or above-average reading abilities read narrative stories with six types of embedded anomalies, with a view toward both accuracy and the use of context to guess the meaning of the anomaly. In the second test, 10 similar students read stories with fewer anomalies; in the third, 20 unskilled readers, aged 9 to 12, read the same stories without anomalies. Results indicated that the necessity of accurate decoding depended on the desired type of comprehension; accuracy appeared to be necessary for atomistic precision but less so for global interpretation. References, graphs, and a sample instrument are included. (DF)
Technical Report No. 118*

EFFECTS OF EMBEDDED ANOMALIES AND ORAL READING ERRORS ON CHILDREN'S UNDERSTANDING OF STORIES

Tom Nicholson
University of Waikato, New Zealand

P. David Pearson
University of Illinois at Urbana-Champaign

Robert Dykstra
University of Minnesota

March 1979

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

The research reported herein was supported in part by the National Institute of Education under Contract No. US-NIE-C-400-76-0116. It was completed while Professor Pearson was on leave from the University of Minnesota.
Embedded Anomalies

Abstract

The basic question asked by this study was, "Do mistakes made while reading interfere with comprehension?" A secondary aim was to clarify the competing views of Gough (1972) who argues that the unskilled reader would not make hypotheses about what words in the text should be, and Goodman (1976), who argues that the generation of tentative hypotheses about meaning are an inevitable part of the reading process. Three experiments were conducted. The first two were simulations in which skilled readers, about 9 years of age, read stories which varied according to error type, difficulty, error rate, text access, and set strength. The design was a $2 \times 6$ factorial using repeated measures. The third experiment was a naturalistic comparison study, using unskilled readers. Overall, the results offered support for both Gough and Goodman, depending on the kind of comprehension tested and the criterion for acceptability of responses. It seems that accuracy is necessary for atomistic precision; less so for global interpretation.
Embedded Anomalies

Effects of Embedded Anomalies and Oral Reading Errors on Children's Understanding of Stories

The purpose of this study was to analyze systematically the relative effect of different types of word misidentification on children's understanding of connected discourse. The basic question asked by the study was: "Do mistakes made while reading interfere with comprehension?"

**Background**

While no one would deny the fact that beginning and unskilled readers often misidentify words, there is a great deal of argument among reading researchers as to whether or not these misidentifications interfere with children's ability to understand written material.

There are those, like Gough (1972, p. 354), who would argue that "since the good reader need not guess [at words] the bad should not." The essence of the Gough view is that word recognition is not influenced or determined by the surrounding words in a story nor by the reader's prior knowledge—the reading process is too rapid for such hypotheses-testing to take place (Cosky & Gough, Note 1; Gough, 1975; Gough & Cosky, 1977). Gough's research suggests that reading is an outside-in or bottom-up process in which the reader processes all of the text data, using it as a base from which meaning is then constructed.

A second view, sometimes referred to as top-down and associated with Goodman (1976a; Goodman & Burke, 1973), implies that misidentification of words need not necessarily interfere with the understanding process.
Goodman (1976b, p. 491) argues that the skilled reader uses "the least amount of information possible to make the best guess possible." In other words, reading is regarded as an inside-out process in which the reader makes tentative hypotheses about the possible meaning of the text, using the graphic array to confirm or revise hypotheses. The disruptive effect of errors, or miscues, depends upon the degree to which they match the tentative meaning hypotheses under consideration by the reader. Those at odds with hypotheses will either be rejected or cause the reader to re-evaluate hypotheses, resulting, perhaps, in some rereading of the graphic array. Those consistent with hypotheses are likely to be accepted at face value; current hypotheses may even be strengthened, and top-down processing is likely to continue.

Related Research

The data base for either point of view, however, is inconclusive. On the one hand, it has been shown that semantically sensible mismatches between text and oral utterance are often accompanied by high levels of comprehension (Goodman & Burke, 1973; Thomas, 1975; Recht, 1976), thus supporting the Goodman view. On the other hand, children often make numerous mismatches that are not sensible yet still are able to exhibit adequate understanding (Biemiller, 1970; Menosky, 1971).

The inconsistency of these findings seems partly due to the influence of uncontrolled variables. First, there is evidence to suggest that the semantic effects of "errors" of different types and from different form classes will vary considerably (Louthan, 1965; Spring, 1976; Weaver &
Bickley, Note 2). Second, results may well depend on how comprehension is measured; whether the probes rely primarily upon textual information or background knowledge (Tuinman & Farr, 1972; Cofer, 1973). Third, results may depend on conceptual difficulty; stories constructed with words and themes familiar to readers may tolerate a higher incidence of oral reading mismatches than those constructed with unfamiliar words and themes.

One way of investigating the above issues would be to ask a group of children who vary in age, ability, and experience to read materials which differ in familiarity, difficulty, and interest. Oral reading errors could be recorded and the degree to which different types of "error" interfere with comprehension could be measured. Yet such research would not be easy to conduct because of the fact that the experimenter is to some extent under the control of (perhaps at the mercy of) the subjects. The experimenter would have to wait for an "error" to occur and then, on the spot, develop some probe for assessing comprehension of the particular text segment in which the "error" occurred. Standard experimental criteria like reliability, replicability, objectivity, and comparability of treatment across subjects would be difficult to achieve because of the fact that the number of errors, their semantic appropriateness, and their form class would vary from subject to subject.

An alternative (though less ecologically valid and hence less satisfying) procedure is to create an experimental simulation of word identification "errors" by embedding anomalous words in the texts, thereby forcing readers to use story context to make sense of certain kinds of simulated
misidentifications. By using a simulation design, therefore, the quality, form class, and number of errors can be predetermined; in other words, it is possible to control some of the complexity which may have confounded the results of previous research.

We decided, in the present study, to conduct a simulation experiment because of the advantage it offered in precision of measurement, and then to follow it up with a naturalistic experiment in order to assess the extent to which the simulated results were anchored in "real world" reading behavior.

Overview

The study consisted of three experiments. The major experiment was a simulation designed to provide systematic control of error types and factors associated with them; the second experiment was a follow-up investigation of the results for the set strength factor in the simulation; the third experiment was a naturalistic follow-up to determine the extent to which the simulation results characterized actual reading behavior.

Experiment 1: Simulation

The purpose of the simulation was to analyze the relative effects of different error types on understanding by simulating the reading environment faced by the unskilled reader who has to answer comprehension questions. It was assumed that in trying to understand a story, the unskilled reader is not only faced with insufficient text data (caused by failing to respond at all to certain words) but anomalous data as well (caused by,
responding with a substitution). By implanting simulated "errors," or anomalies, within stories read by skilled readers, we hoped that the resulting text interpretation would resemble the kind of story which poor readers create when they are unable to decode accurately all the words in the text. We used the simulation design in order to control those factors which are influential in the actual reading situation but which are difficult to control in a naturalistic design. The simulation was complex, but necessarily so because of the need to approximate what happens in actual reading situations where factors such as error type, their rate of occurrence, story difficulty, text accessibility, and repetition (set strength) seem to interact to influence children's understanding of narrative stories.

Method

Subjects. Eighty children (41 girls, 39 boys), all about 9 years of age, at or above grade level in reading, were selected from a middle class suburb of the Twin Cities, Minnesota.

Task. Subjects read and answered questions about six stories. Each story was transformed so that it contained simulated errors of a particular type. It was assumed that the proficient readers could decode the simulated errors accurately.

Materials. The stimulus materials consisted of 12 short narrative stories (each 100 words in length), all selected from basal reading material. Passage difficulty was varied by including six stories at grade 4 level and six stories at grades 5-6 level (Dale & Chall, 1948; Fry, 1963).
Stories were transformed by embedding different types of simulated errors or anomalies (called "simulates" to avoid the connotatively biased terms "error" and "miscue") in place of randomly selected nominals. There were six simulate conditions used, operationally defined according to their semantic and visual relation to the target word:

- **CORR** - the target word originally in the text (e.g., dragon)
- **SRVU** - semantically related, visually unrelated (e.g., monster)
- **SUVR** - semantically unrelated, visually related (e.g., doctor)
- **SUVU** - semantically/visually unrelated (e.g., rabbit)
- **NONE** - a non-response (simulated by a blank space in the text)
- **MIX** - mixed simulate type, used to simulate the naturalistic situation: that is, all simulate types except CORR were included.

Each story was transformed according to each simulate type, so that there were six versions of each story (see Figure 1 for the matrix of simulate types for one experimental story).

---

Insert Figure 1 about here

---

Rate of simulate substitution was varied by replacing either 15 percent or 6 percent of the story's words with simulates with the restriction that only nominals could be replaced with simulates (nominals were selected to maximize the anomalous effect of the simulates). The 6 percent error
rate was selected to correspond to the commonly accepted 95 percent criterion for instructional level used by most informal reading inventories. In other words, at 6 percent error rate, according to the conventional wisdom, most children should be able to cope with the text at hand. The 15 percent error rate was selected to approximate a genuinely frustrating situation. In general, the conventional wisdom regarding informal reading analysis suggests that error rates above 5-10 percent correspond to a frustration level for students.

Set strength was varied either by repeating correct forms of target words in the story title and later in the text (high set) or else just once in a story and not in the title (low set). The set strength factor was included to simulate the influence of conflicting bottom-up data. In short, what happens to a reader when he or she misreads a word in one sentence but reads it correctly elsewhere?

Text access during the comprehension probe phase of two experiments was varied by allowing half of the subjects to look back at the text while answering questions while half the subjects were denied access to the text. Text access was included to evaluate the durability of the anomalous information embedded in the text. That is, would students be more likely to use the embedded simulates when they had the opportunity to look back at it while responding to a comprehension probe?

Dependent measures. Explicitly dependent comprehension was measured by using a cloze-type measure (see Figure 1). A cloze-type rather than a WH-type format was used because a pilot study revealed that WH-questions
provided additional text cues (intended to reduce ambiguity) and also tended to cue the form class of the target word. For example, a what or who question strongly suggests a nominal response while a which question suggests an adjectival response. Inferential comprehension was measured by using a multiple choice format for implicitly dependent (text dependent) questions.

These three question types (after Pearson & Johnson, 1978) represent decreasing dependency on textual information. In the explicitly dependent category, the question and the answer are derivable from the text, and the semantic relationship between question and answer is heavily cued by the syntactic structure of the sentence from which the question is derived, as in examples (1) and (2).

(1) The ghost chased the bear.
(2a) Who chased the bear?
(2b) The _______ chased the bear.

Implicitly dependent comprehension corresponds to Pearson and Johnson's textually implicit category, in which both question and answer are derivable from the text but the relationship between them is not well cued by the syntax of the text, as in examples (3) and (4).

(3) The ghost chased the bear. The bear ran faster.
(4) Why did the bear run faster?

Scriptal comprehension corresponds to Pearson and Johnson's scriptally dependent comprehension, in which the question is derivable from the text but the only plausible answer must come from a reader's prior experience or scriptal (after Schank, 1972) knowledge, as in (5) and (6).
The ghost chased the bear.

Why did the ghost chase the bear?

These measures enabled an assessment of comprehension at an atomistic as well as at a global level of understanding.

Design and randomization procedures. The research design was a $4 \times 6$ factorial. The between-subjects factors were passage difficulty, rate of simulate substitution, and text access. The repeated measures factors were simulate type (6), level, and set strength. Random procedures in the development of materials and assignment of treatments were used wherever possible.

Procedure. The children were shown, as an example, a short story in which an anomalous word occurred. The story was discussed and a test question given. Children were told to hypothesize what the anomalous word in the text should mean—in other words, use context to make the best guess possible about the real meaning of the word. After a discussion of the sample story, children were given the experimental set of stories and questions, and told to ask for help with any words they found difficult to read. In brief, children were encouraged to read accurately but to try to make sense of the stories they read.

Administration. The data were collected, using standardized test protocols, by eight graduate students, all testing taking place at the same time.

Analysis. The cloze responses were scored according to a 9-point semantic appropriateness scale (Figure 2) and analyzed according to two
dichotomous criteria--strict criterion, where the exact word was required for correctness; broad criterion, where response types 5-9 on the semantic appropriateness scale were scored as correct. Responses were also scored according to whether they directly matched the corresponding text simulate. Finally, inferential comprehension was analyzed as a fourth dependent measure, with two levels: implicitly dependent and scriptal. Statistical procedures involved 5- and 4-way ANOVAS for major and follow-up analyses.

Results

Simulate type. When strict criterion scores were analyzed, there was a main effect for simulate type, $F(5, 360) = 93.67$, $p < .01$ ($\bar{X}_{CORR} = .7792$, $\bar{X}_{SRVU} = .2292$, $\bar{X}_{SRVR} = .3083$, $\bar{X}_{SUVU} = .2438$, $\bar{X}_{NONE} = .3958$, $\bar{X}_{MIX} = .2708$). Post hoc comparisons, using Student-Newman-Keuls (SNK) procedures, showed that CORR and NONE were significantly superior to all other simulate types and different from each other. When broad criterion scores were analyzed, there was also a main effect for simulate type, $F(5, 360) = 60.63$, $p < .01$ ($\bar{X}_{CORR} = .8896$, $\bar{X}_{SRVU} = .7458$, $\bar{X}_{SRVR} = .4563$, $\bar{X}_{SUVU} = .4188$, $\bar{X}_{NONE} = .6379$, $\bar{X}_{MIX} = .5354$). Follow-up comparisons, using SNK procedures, showed that CORR, SRVU, and NONE, respectively, were superior to SUVU, SUVR, and MIX, and significantly different from each other. The results for inferential comprehension showed no main effect for simulate type, $F(5, 360) = 1.70$, $p > .01$. Nor was there a significant difference between implicitly
dependent and scriptal questions, $F(1,72) = 0.77, p > .01$ ($\overline{X}_{ID} = .7188$, $\overline{X}_{SC} = .6917$). In brief, as can be seen from Figure 3, the overall results indicated that the effects of simulate type depend on both the comprehension task and the criterion used to assess responses. The simulated errors were most disruptive when accuracy was required; less so when broad understanding was required; not at all when inferential understanding was required.

Set strength. When strict criterion scores were analyzed, there was a main effect for set strength, $F(1,72) = 37.44, p < .01$ ($\overline{X}_{HS} = .4139$, $\overline{X}_{LS} = .3285$). The results for broad criterion scoring showed that set strength interacted with story difficulty, $F(1,72) = 13.4, p < .01$ ($\overline{X}_{HD} = .6500$, $\overline{X}_{LD} = .5542$, $\overline{X}_{HE} = .6181$, $\overline{X}_{LE} = .6633$). These results suggest that it is presumably more useful for the reader to be aware of important target concepts (as in a title) before reading the story when the story includes relatively difficult content.

Passage difficulty. Passage difficulty interacted with set strength, text access, and rate of simulate substitution. The nature of the interactions are explained under the results for the other factors.

Text access. Text access interacted with story difficulty, $F(1,72) = 13.10, p < .01$ ($\overline{X}_{AE} = .4306$, $\overline{X}_{NAE} = .3514$, $\overline{X}_{AD} = .2972$, $\overline{X}_{NAD} = .4056$) for strict criterion scoring, and for broad criterion scoring, $F(1,72) = 13.25$, 14
Embedded Anomalies

13

\( P < .01 (\overline{X}_{AE} = .6722, \overline{X}_{NAE} = .5792, \overline{X}_{AD} = .5473, \overline{X}_{NAD} = .6569) \). In order to analyze the interactions more carefully, separate ANOVAs for easy and difficult story results were calculated. Results showed no main effect for text access in difficult stories in favor of no-access for both strict, \( F(1,36) = 8.99, p < .01 \), and broad, \( F(1,26) = 9.27, p < .01 \), criterion scoring. Preventing the reader from looking back at difficult anomalous material facilitated explicitly dependent comprehension. Results also showed, when matching responses (the student gave, as an answer, the exact simulate in the text) were analyzed, a main effect for text access, \( F(1,72) = 22.12, p < .01 \) (\( \overline{X}_{AM} = .3927, \overline{X}_{NAM} = .2021 \)). What seems to happen is either a "potency" effect (as Thorndike, 1917, would have described it) or else the reader is "pattern matching" (Pearson, 1978) the question with the text. The simulates interfere with comprehension when text access is possible. When access to the text is blocked, readers seem to revert to prior knowledge to answer questions.

Rate of simulate substitution. There was no main effect either for strict criterion scoring, \( F(1,72) = 0.52, p > .01 \) or for broad criterion scoring, \( F(1,72) = 3.43, p > .01 \). The results suggest that rate of error may not disrupt comprehension unduly after a certain point, keeping in mind that for accurate understanding, a 6 percent nominal error rate is already highly disruptive. One must remember, however, that relative to the 6 percent condition, all the 15 percent condition did was to destroy the context surrounding the probed constituents. The very same constituents were probed in both the 6 and 15 percent conditions. Viewed from another
perspective, these results suggest that when comprehension is measured atomistically, anomalizing the surrounding context by an additional 9 percent has no effect on children's ability to respond sensibly to such probes. It suggests the possibility that this type of comprehension is not influenced greatly by surrounding context.

Experiment 2: Set Strength Follow-Up

The purpose of the follow-up was to determine whether a single repetition of the correct word (low set) was any more useful than no repetition at all. In the low set condition, the behavior simulated was that in which the reader has a chance to correct a previously misidentified word. When this happens it is sometimes inferred (Goodman, 1976a; Clay, 1968) that the reader has rejected, in memory, the previous error. If this is the case, then we should expect that the single repetition of the correct word after the initial error would be more helpful than no repetition at all.

Method

Subjects. Ten subjects, all about 9 years of age and above average in reading ability, were selected.

Task. Subject read eight stories (four easy, four difficult) and answered comprehension questions.

Materials. There were four versions of each story: CORR, SRVU, SUVR, and NONE. Set strength was varied so that of the six simulates two words were not repeated (no set), two were repeated correctly once later in the story (low set), and two were present twice in their correct form, once later in the story and once in the title (high set).
Analysis and design. The design was a $4 \times 2 \times 3$ factorial, with repeated measures on all factors. Responses were scored by strict and broad criteria and analyzed by using 3-way ANOVA procedures.

Results and Discussion

The results for strict criterion scoring showed a significant main effect for set strength, $F(2,216) = 6.00, p < .01 (\bar{X}_N = .3626, \bar{X} = .4063, \bar{X}_H = .5188)$. Follow-up comparisons using Duncan's new multiple-range test (Winer, 1971) showed that high set was superior to low ($p < .01$) with no difference between the no and low set means ($p > .05$).

The results for broad criterion scoring showed no main effect for set strength, $F(2,216) = 2.21, p > .05 (\bar{X}_N = .7063, \bar{X}_L = .7188, \bar{X}_H = .7938)$, although the results were in the same direction as for strict criterion scoring.

Put briefly, the results indicated that for strict scoring the effect of a single repetition of the correct word did not differ from no repetition at all. The broad criterion scoring indicated that students were remarkably capable of getting the semantic sense of the story even when there was no textual evidence to contradict an embedded anomaly.

Experiment 3: Naturalistic Follow-Up

The purpose of the naturalistic experiment was to find out the extent to which the simulation results characterized actual reading behavior.
Method

Subjects. Twenty subjects (14 boys, 6 girls), all unskilled readers, aged 9-12, were selected.

Materials. Eight of the simulation stories in their correct versions (four easy, four difficult) were used.

Procedure. Four treatments (the four combinations of easy or difficult and access or no access) were randomly assigned to subjects. Subjects were tested individually. Each subject read each story aloud to the experimenter. The experimenter recorded all oral reading errors but paid particular attention to errors which were made on those words which had been manipulated as simulates in the earlier experiments. After reading a story, each subject answered the six cloze comprehension probes and the two inferential probes used in the previous experiments.

Scoring. Each time a subject made an error on one of the key words (i.e., those words for which comprehension probes had been developed), the error was categorized according to the four types of errors simulated in the previous experiments—SRVU, SUVR, SUVU, NONE. Then the subject's answer to the probe was scored correct or incorrect according to both strict and broad scoring criteria. In addition, the two inferential questions for each story were scored as correct or incorrect.

Analysis. Data were analyzed descriptively. The results are reported in empirical probabilities. For example, a p = .06 means that in 6% of the cases in which students made a particular type of error, they were able to answer the cloze question probe correctly. In addition, the number (N) of
such errors on which the proportion was calculated is provided. Results are summed, not averaged, across subjects.

Results and Discussion

Only the main results are reported. For the strict scoring criterion, errors proved highly disruptive (p = .06, N = 76). They were less disruptive (p = .38, N = 76) when the broad scoring criterion was applied.

There were too few SRVU and SUVU errors on keywords to merit analysis. There were 24 NONE errors and 52 SUVR errors made on keywords. Using the strict criterion, NONE errors elicited more correct answers (p = .13) than SUVR errors (p = .04). However, the broad scoring criterion suggested a reversal, with SUVR errors allowing more correct answers (p = .43) than NONE errors (p = .33).

The effect of text access is somewhat puzzling. When students were allowed to look back at the stories to answer questions, they responded with the SUVR error they had made while reading orally (p = .43, N = 24) more often than when they were not allowed to look back (p = .00, N = 28). Summed across SUVR and NONE errors, according to the broad criterion, no access was somewhat more favorable (p = .40, N = 42) than text access (p = .36, N = 34). There were virtually no differences between access conditions when the strict scoring criterion was applied. Ironically, looking back at the text seemed to strengthen the probability of responding with the error made during oral reading and to decrease the likelihood of a semantically acceptable response. These results are reminiscent of the access effect in the simulation experiment for difficult stories, when
no-access was more helpful than access. Taken in concert, these results seem to suggest that when reading becomes difficult, concentrating on textual features is less likely to result in acceptable semantic interpretation than is relying on whatever knowledge structures have been instantiated in the process of trying to read the text, recognizing, of course, that under no-access conditions, one has no recourse but to rely on knowledge structures.

As in the simulation study, errors had little effect on the more global comprehension required in the inferential probes. As an added analysis, each oral reading of each story by each student was classified as exhibiting a high rate of oral reading errors or a low rate of oral reading errors. The inferential probes were examined as a function of error rate. Error rate affected the comprehension of implicitly dependent probes, with low error rate instances ($p = .81$) eliciting better comprehension than high error rate instances ($p = .49$). In contrast, error rate did not affect scriptal comprehension ($p_{H} = .70$, $p_{L} = .70$). This difference seems reasonable in light of the realization that implicitly dependent comprehension requires more attention to text than does scriptal comprehension.

It would strain even an ardent believer's imagination to suggest that the naturalistic results completely corroborate the results of the simulation experiment. Yet the results, for the most part, are in the right direction.
General Discussion

The most consistent finding across all three studies is that the necessity of accurate decoding depends upon the type of comprehension one considers important. If one wants precise atomistic comprehension, then accurate decoding seems to be a requisite behavior. Alternatively, one wants global interpretation, accurate decoding seems relatively unimportant. Hence, the incipient tension between the Gough and Goodman viewpoints seems resolvable; it all depends on what goals are set for comprehension.

There is a natural temptation to prefer global interpretation over precise atomistic comprehension. And in most situations global interpretation is a more desirable form of comprehension. Yet, there will surely be instances, particularly in instructional settings, when it is important "to get the facts straight." In such instances, gross semantic acceptability will simply not suffice. No matter how sincere the reader's attempt to impose meaning onto a text that seems to defy interpretation, he or she will quite often answer detail questions incorrectly.

Furthermore, semantically acceptable errors, because they seem so reasonable, are likely to disrupt precise atomistic comprehension more than failures to respond to overtly or semantically unacceptable errors: the child who reads "giant" as "gorilla" is more likely to maintain that interpretation than a child who reads "giant" as "wall."

The effect of set strength suggests that the self-correction hypothesis is difficult to support empirically. Students were no more likely to reject an anomalous simulate when they later encountered a correct form of the word...
than when they did not. On the other hand, if the correct form is made salient (i.e., as a part of a title) prior to encountering the anomalous simulate, then students are more likely to overcome the effect of the simulate. We do not want to suggest that students do not engage in self-correction. There is ample experimental and clinical evidence that they do (Goodman & Burke, 1973). All we are bringing into question is the ubiquity of the behavior at the same time that we suggest that there are many instances in which students may not recover from misidentifications.

The failure of rate of simulate substitution to demonstrate an effect suggests that atomistic comprehension is so atomistic that it is unaffected by anomalizing a substantial proportion of story context. On the other hand, comprehension probes requiring a student to relate two text segments were affected by error rate in the naturalistic study. Finally, scriptal comprehension was unaffected by error rate in any of the studies, suggesting its primary reliance on prior knowledge.

The effects of text access while answering question probes appears to be counter-intuitive. In many instances, students were better off when they could not look back at the text. Recall that in the simulation experiment this advantage occurred only for difficult stories. It is almost as though the lack of familiar content in the difficult stories disposed those students who had access to the text to trust their prior knowledge structures very little—a poor decision in view of the fact that those who had to rely on instantiated knowledge structures exhibited better comprehension, particularly when a broad scoring criterion was employed. In this matter, the
results of the naturalistic follow-up study provided reasonable corroboration. It may be that the problem with text access is that students repeat their bad habits, thus strengthening the inappropriate response. Certainly, the high proportion of repeated SUVR responses when text access was permitted supports such an interpretation.

Taken as a whole, these studies suggest that teachers need to have a clear grasp of what they want students to gain from reading a selection. Having made that decision, teachers can encourage differential processing strategies as a function of the comprehension goals they help students set. Further, they can suggest to students that when they read on their own, different comprehension goals will dictate different strategies for interacting with text.

In terms of underlying theories of the reading process, the data are appropriately ambiguous, suggesting that both the Goodman and Gough models must be precise about the type of comprehension under consideration. It is, however, perhaps unfair to single out these two models, since few, if any, models of reading are very specific about the nature of the task demands imposed during encounters with text. Indeed, recent thinking and research suggest the need to move toward a model of reading in which the use of text data and prior knowledge structures in story understanding varies according to the complexity of the comprehension task, the familiarity of the text, and the level of understanding required of the reader (Pearson & Nicholson, Note 3; Pearson, Note 4; Nicholson & Imlach, Note 5). Only an interactive model like that Rumelhart (1977) has developed seems
capable of gracefully incorporating such concerns. We admit that Rumelhart has not spoken to these issues; however, we believe that such variables will be easier to explain in an interactive than in a top-down or bottom-up framework.

Limitations

The use of a simulation design had certain disadvantages. Generalizability was lost in simulating rather than measuring naturally occurring behaviors. What was gained, however, was precision—precision in estimating the parameters under which certain types of oral reading behavior do and do not interfere with comprehension. In addition, the naturalistic follow-up experiment enabled an investigation of the extent to which the simulation findings were anchored in actual reading behavior. Nevertheless, further replication research in actual reading situations is needed before firm conclusions can be drawn.

Future Research

An interesting future development in studying the effects of errors on understanding would be to focus on a variety of comprehension tasks, such as the ability to retell, summarize, and paraphrase the event structures of stories (Bower, 1976; Thorndyke, 1977; Mandler & Johnson, 1977) as well as children's question-answering ability. Future research is also needed to clarify the effects of errors made at the proposition or paragraph level rather than at the word level.
Concluding Statement

The results of the study strongly suggest that future research into the semantic effects of oral reading errors must take into account the way comprehension is assessed. It seems clear that the necessity for precise decoding depends upon whether comprehension is assessed in its atomistic or global aspect: precision is important for atomistic detailed comprehension, less so for global interpretation.
Reference Notes


References


Biemiller, A. The development of the use of graphic and contextual information as children learn to read. Reading Research Quarterly, 1970, 6, 75-96.


Goodman, K. S. Reading: A psycholinguistic guessing game. In H. Singer & R. Ruddell (Eds.), Theoretical models and processes of reading (2nd ed.). Newark, Del.: International Reading Association, 1976. (a)

Goodman, K. S. Behind the eye: What happens in reading. In H. Singer & R. Ruddell (Eds.), Theoretical models and processes of reading (2nd ed.). Newark, Del.: International Reading Association, 1976. (b)


Footnotes

1 Later, when we introduce independent variables, we will use the term "simulate" to describe our embedded anomalies. We have tried to avoid, where possible, the indiscriminate use of the term "miscue" or even "simulated miscue," because of the special meaning assigned to those terms by particular researchers. We also recognize the fact that Goodman (1976a) did not mean for the term "miscue" to be applied exclusively to words, although he does allow that there is at least a surface level similarity between what others have called errors and what he calls miscues (1976a, p. 499). And occasionally, he does use the term error (1976b, p. 492), although either term, for Goodman, involves the stipulation that the reader committed them in his or her constant search for meaning and that the reader is capable of self-correcting them when the meaning he or she imposed on the text suggests correction. We would doubt, however, that even Goodman would deny that, at the surface level at least, what he calls miscues "look like" what others have called oral reading errors or that they "look like" mismatches between words in the text and words readers utter while reading.

2 The syntactic relation of simulates to the target words was not included because of the fact that most word substitutions are of the same form class (Clay, 1968; Weber, 1968; Goodman & Burke, 1973).

3 The term "scriptal" is derived from Schank's (1972) and Abelson's (1973) notion of "script" as a representation of prior knowledge in memory.

4 At first glance such a finding might seem to invalidate the original set of simulates chosen for the simulation experiment. However, one must
remember that we analyzed only those words which had been used as simulate positions in Experiment 1. In fact, what is remarkable is the high incidence of visually similar errors and the low incidence of semantically related errors among this set of nominals.
Figure Captions

Figure 1. An actual story used in the study (easy in difficulty level), with comprehension probes used to test understanding. (Asterisks indicate those nominals randomly selected for the six percent rate condition. Words circled indicated high set strength repetitions.)

Figure 2. A classification scheme for scoring the semantic appropriateness of questions testing explicitly dependent comprehension.

Figure 3. The relative effects of different simulate types on explicitly dependent comprehension when scored according to strict and broad criteria, and on inferential comprehension.
Once there was a farmer. He had lots of fine animals on his farm. But there was a ghost in the house. It would often walk around in the night and take the sheets off everyone. Sometimes it knocked on the door. When the sleepy farmer got up, no one was there. Sometimes it rolled a bottle down the stairs. It made a great noise. One night the ghost rattled all the pots and scared the animals. It also liked to blow the smoke back down the chimney. Then no one could light a fire. The smoke would go everywhere. The fire would go out.

Matrix of Simulate Types to Replace Target Nominals

<table>
<thead>
<tr>
<th>CORR</th>
<th>SVRU</th>
<th>SUVR</th>
<th>SUVT</th>
<th>NONE</th>
<th>MIX</th>
</tr>
</thead>
<tbody>
<tr>
<td>farmer</td>
<td>grower</td>
<td>factory</td>
<td>ship</td>
<td></td>
<td>grower</td>
</tr>
<tr>
<td>animals</td>
<td>livestock</td>
<td>ankles</td>
<td>ladders</td>
<td></td>
<td>ankles</td>
</tr>
<tr>
<td>farm</td>
<td>land</td>
<td>frame</td>
<td>knife</td>
<td></td>
<td>land</td>
</tr>
<tr>
<td>ghost</td>
<td>demon</td>
<td>glove</td>
<td>rope</td>
<td></td>
<td>rope</td>
</tr>
<tr>
<td>house</td>
<td>cabin</td>
<td>hose</td>
<td>river</td>
<td></td>
<td>host</td>
</tr>
<tr>
<td>night</td>
<td>evening</td>
<td>nest</td>
<td>fig</td>
<td></td>
<td>nest</td>
</tr>
<tr>
<td>sheets</td>
<td>cloth</td>
<td>shells</td>
<td>lights</td>
<td></td>
<td>shells</td>
</tr>
<tr>
<td>door</td>
<td>gate</td>
<td>doll</td>
<td>spot</td>
<td></td>
<td>gate</td>
</tr>
<tr>
<td>bottle</td>
<td>glass</td>
<td>ball</td>
<td>dust</td>
<td></td>
<td>dust</td>
</tr>
<tr>
<td>stairs</td>
<td>steps</td>
<td>stones</td>
<td>collar</td>
<td></td>
<td>steps</td>
</tr>
<tr>
<td>noise</td>
<td>sound</td>
<td>nose</td>
<td>paint</td>
<td></td>
<td>nose</td>
</tr>
<tr>
<td>pots</td>
<td>bowls</td>
<td>pets</td>
<td>faces</td>
<td></td>
<td>faces</td>
</tr>
<tr>
<td>smoke</td>
<td>fumes</td>
<td>snail</td>
<td>guns</td>
<td></td>
<td>snail</td>
</tr>
<tr>
<td>chimney</td>
<td>pipe</td>
<td>Chicken</td>
<td>forest</td>
<td></td>
<td>forest</td>
</tr>
<tr>
<td>fire</td>
<td>blaze</td>
<td>fish</td>
<td>pie</td>
<td></td>
<td>pie</td>
</tr>
</tbody>
</table>

Comprehension Probes

Explicitly Dependent:
1. But there was a ______ in the house.
2. It also liked to blow the ______ back down.
3. It would often walk around in the ______
4. But no one could light a ______
5. He had lots of fine ______
6. Once there was a ______

Inferential
7. Why did the farmer get up?
   A. He heard the ghost.
   B. He could no sleep.
   C. He wanted to lock the door.
   D. He wanted to go down the stairs.

8. Why did the ghost roll the bottle down the stairs?
Once there was a king. He lived in an old castle. There was no carpet. The walls were made of stone. It was cold. Sometimes he would wear a blanket.

Outside there lived a dragon. Every night he sat in his cave on the hill and roared. The king could not get any sleep. Finally the knights were sent out. But he frightened them all away from the cave, except one. The dragon started crying: "I have no friends." The knight said: "I'll be your friend."

Now he lives in the castle. The fire comes out of his nose and keeps everyone warm.

Target Structure: Outside there lived a ________

Classification Scheme:

Response to Target Stimulus

Visually & Semantically Identical

- Superordinate
- Coordinate
- Subordinate

- Dragon
- Monster
- Dinosaur
- Claw
- Giant
- Doorbell
- Poster
- King
No. 1: Durkin, D. Comprehension Instruction—Where are You?, October 1977. (ERIC Document Reproduction Service No. ED 146 566, 14p., HC-$1.67, MF-$0.83)

No. 2: Asher, S. R. Sex Differences in Reading Achievement, October 1977. (ERIC Document Reproduction Service No. ED 145 567, 30p., HC-$2.00, MF-$0.83)


No. 4: Jenkins, J. R., & Pany, D. Teaching Reading Comprehension in the Middle Grades, January 1978. (ERIC Document Reproduction Service No. ED 151 756, 36p., HC-$2.06, MF-$0.83)

No. 5: Bruce, B. What Makes a Good Story?, June 1978. (ERIC Document Reproduction Service No. ED 158 222, 16p., HC-$1.67, MF-$0.83)


No. 4: Alessi, S. M.; Anderson, T. H., & Biddle, W. B. *Hardware and Software Considerations in Computer Based Course Management*, November 1975. (ERIC Document Reproduction Service No. ED 134 928, 21p., HC-$1.67, MF-$0.83)


No. 15: Schwartz, R. M. *Strategic Processes in Beginning Reading*, November 1976. (ERIC Document Reproduction Service No. ED 134 938, 43p., HC-$2.06, MF-$0.83)

No. 16: Jenkins, J. R., & Pany, D. *Curriculum Biases in Reading Achievement Tests*, November 1976. (ERIC Document Reproduction Service No. ED 134 939, 24p., HC-$1.67, MF-$0.83)


No. 20: Kleiman, G. M. *The Effect of Previous Context on Reading Individual Words*, February 1977. (ERIC Document Reproduction Service No. ED 134 941, 76p., HC-$4.67, MF-$0.83)


No. 105: Ortony, A. Beyond Literal Similarity, October 1978.

No. 106: Durkin, D. What Classroom Observations Reveal about Reading Comprehension Instruction, October 1978. (ERIC Document Reproduction Service No. ED 162 259, 94p., HC-$4.67, MF-$0.83)


