DOCUMENT RESUME

ED 215 307 CS 006 601

AUTHOR Pearson, P. David
TITLE A Context for Instructional Research on Reading Comprehension.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
REPORT NO CSR-TR-230
PUB DATE Feb 82
CONTRACT 400-76-0116
NOTE 34p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Classroom Research; *Cognitive Processes; Elementary Secondary Education; Epistemology; *Literature Reviews; Reading Achievement; *Reading Comprehension; Reading Instruction; *Reading Research; Research Methodology; *Teaching Methods; Time on Task
IDENTIFIERS Direct Instruction; *Reading Strategies

ABSTRACT

Cognitive research of the 1970s has shown that both content factors (topical world knowledge and knowledge about textual organization) and process factors (attention, encoding, inference, retrieval, and executive monitoring) influence comprehension. Classroom research during the same decade has shown that the greater the proportion of time students spend on a task, the better their performance on the task; content covered tends to be positively related to achievement; that error rate seems to add a significant amount of power in predicting achievement above and beyond engagement and content covered; and that group instruction (particularly small group instruction) is consistently associated with positive gains in achievement. The research surveying current practices for teaching reading comprehension, however, is limited to two studies by D. Durkin that showed virtually no direct instruction in comprehension. Other recent studies have evaluated the effects of direct explicit attempts to help students develop heuristic strategies (if not rules) for dealing with a range of comprehension tasks typically required in schools. These data suggest comprehension skills can be taught. Future research should focus on explicit attempts to help students develop independent strategies for coping with the kinds of comprehension problems they are asked to solve in school. Current knowledge of basic comprehension processes and instruction, while not complete, is sufficient to allow its application to issues of reading comprehension instruction. (HOD)

CENTER FOR THE STUDY OF READING

Technical Report No. 230

A CONTEXT FOR INSTRUCTIONAL RESEARCH ON READING COMPREHENSION

P. David Pearson

University of Illinois at Urbana-Champaign

February 1982

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. Resources in the form of graduate assistantships were also provided by Ginn & Company, Lexington, Massachusetts.

EDITORIAL BOARD

A Context for Instructional Research on Reading Comprehension

When the history of reading research is written for this century, the decade of the seventies will be regarded with irony. During the seventies, our knowledge of the basic cognitive processes involved in reading comprehension as well as our knowledge about basic instructional processes grew dramatically. Nonetheless, our knowledge about teaching reading comprehension advanced very little, if at all. This ironic state of affairs can be pardoned, perhaps, on grounds that we had to learn about the basic processes of comprehension and the basic processes of instruction before we could combine knowledge about these two areas in order to make any advances in applied knowledge about teaching reading comprehension.

The time has come to begin a vigorous program of research that directly addresses the issue of how we can improve the reading comprehension abilities of our students in our schools. Our knowledge of basic processes, while not complete, is sufficient to allow us to begin to apply knowledge about comprehension and instruction to issues of reading comprehension instruction. Moreover, even if we did not have the benefit of basic process knowledge, we should still begin the applied effort. Literacy is too important a concern to allow us the luxury of waiting for further advances in basic research. Besides, the argument for waiting reveals an elitist fallacy about the relationship between basic and applied research. There is no reason why applied research cannot and should not inform basic research in the same measure that basic research informs applied.

In this paper I will summarize briefly important conclusions from research on basic cognitive processes involved in reading comprehension. Second, I will do the same for research on classroom instruction. Third, I will discuss what we have learned about how reading comprehension is taught (or is not taught) in today's schools. Fourth, I will discuss the few experimental studies that have been conducted in which experimenters have tried to intervene in the ecology of the school in order to improve students' reading comprehension. Finally, I will speculate about promising directions that such research might take.

Basic Cognitive Processes in Reading Comprehension

The first thing to note about the cognitively oriented research of the 1970s is that it was not so much directed toward reading comprehension as it was toward understanding how information of any sort, including information represented by graphic symbols on a page, is stored and processed. In other words, the research has been as much about attention, encoding, inference, memory storage, and retrieval as it has been about reading comprehension. This is as it should be. It would be counter-intuitive and counterproductive to focus exclusively on reading comprehension, as if separate mechanisms and separate processes were necessary for processing print as opposed to auditory or other visual information. A unified theory of cognitive processing seems a more reasonable possibility than does a set of separate theories.

The most basic conclusion of this research is that reading, and especially reading comprehension, is a complex interactive process (Rumelhart, 1977; Stanovich, 1980)--one in which a reader varies his focus along a continuum from primarily text-based processing (concentration on

getting the author's message straight) to primarily reader-based processing (concentration on predicting what the author's message will likely be). This variation in focus is determined by a number of intertwined factors: reader purpose (What do I have to do with this information once I've read it?), familiarity (How much do I already know about the topic addressed in the text?), interest and motivation (How much do I care about learning this subject?), and discourse type and complexity (How much do I already know about the conventions involved in this particular mode of discourse?).

That the type of processing in which a reader engages is determined by so many factors is at once a curse and a blessing. The curse is that this inherent complexity may make it difficult for us to understand, let alone improve, reading comprehension processes of students. The blessing is that with so many factors involved, the likelihood increases that we will find a small subset of factors--or even one factor--that we can manipulate systematically with the result of improved comprehension. Our hope, therefore, may reside in being able to select those factors most amenable to improvement through instruction.

A second conclusion to be drawn from basic research in cognition is that both content and process factors are implicated in reading comprehension. Content factors are the knowledge structures residing in our long-term semantic memory that determine how well we understand and integrate a particular text. They are like what computer scientists call data structures. To put it simply, the more we know about the topic addressed in the text, the greater the likelihood we will understand, integrate and remember the information contained in the text. Such a likelihood has indeed been verified in a number of studies (e.g., Anderson, Reynolds, Schallert, &

Goetz, 1977; Pearson, Hansen, & Gordon, 1979). But there is another type of content, besides knowledge of topic, that influences comprehension--knowledge about the text structure or text genre in which the topical content is embedded. The work on story structures (e.g., Neilsen, 1977; Omanson, in press; Stein & Glenn, 1979; Thorndyke, 1977) and typical rhetorical structures found in expository writing (e.g., Meyer, 1977; Meyer, Brandt, & Bluth, 1980) indicates that familiarity with structure influences comprehension. Neilsen (1977), for example, found that even when topical information (as defined by the characters and activities) was controlled, subjects were better able to recall and recognize information presented in a causally organized structure than they were information presented in a mere sequentially organized structure. Several studies (e.g., Stein & Glenn, 1979; Mandler, 1978; Thorndyke, 1977) have indicated that violations in what might be labelled canonical story form result in a decrement in recall of information. The point, in terms of content, is that both topical and structural content have identifiable influences on comprehension.

Process factors are comparable to what are called control procedures in computer processing. They refer to how data are processed instead of what data are processed. To discuss them in a paragraph separate from content factors may seem to imply that I think they are separate from and independent of content factors. If that implication exists in the reader's mind, it should be regarded as an accident of the conventions of print. I know of no data base that would allow us to determine the independence of content (data) and process (control) factors. Process factors may be but different facets of the same amalgam under consideration when content factors are discussed. The kinds of procedures I have in mind are attention,

encoding, inference, retrieval, as well as executive monitoring of these procedures (what some people refer to as metacognitive processing--knowledge about the procedures or how they are "proceeding").

That these processes undergo developmental improvement seems intuitively obvious. In fact, empirical researchers have indicated such a trend for processes like inference (Paris & Lindauer, 1976; Paris & Upton, 1976), encoding of information into memory (Pichert & Anderson, 1977), retrieval of information from memory (Pichert, 1979), and metacognitive monitoring (Baker & Brown, in press). What is not clear in most of these studies is the factor or factors to which this growth should be attributed--a sheer developmental increase in cognitive capacity, an increase in subjects' world knowledge, instructional history (i.e., schooling), or a growing awareness that the processes are available and ought to be used.

For example, regarding inference, Paris and Lindauer (1976) seem to argue for an awareness of strategy availability, while Trabasso (1981) argues for knowledge changes. Chi (1978) presents evidence favoring a growth in world knowledge as a major determinant of retrieval from memory. Alternatively, recent studies by Hansen (1981) and Gordon (1980) suggest that inference performance increases with direct instruction and/or practice.

The point that can be made to conclude this section is that both content factors (as defined by topical world knowledge and knowledge about textual organization) and process factors (as represented by attention, encoding, inference, retrieval, and executive monitoring) have been shown to influence comprehension.

Classroom Instruction Research

In the past decade researchers have spent a great deal of time in classrooms; observing what goes on in that environment. The general paradigm for the research is based upon the assumption that observation techniques will allow us to identify management, material, design, and verbal interaction patterns that discriminate between successful and unsuccessful classrooms and/or schools. This is typically accomplished by identifying, in advance, successful and unsuccessful schools, teachers, or classrooms. Then, depending upon the degree to which one accepts the tenets of the ethnographic tradition, the researcher conducts controlled (preplanned, systematic, and theoretically determined) or uncontrolled observation (observing as much of the ecology as possible without predetermined scales or protocols). Then, the researcher examines the observational data, looking for factors that discriminate between successful and unsuccessful sites.

The logic of this paradigm is similar to the good/poor reader paradigm used in descriptive reading research: Give similar tasks to readers with widely different reading ability, and look for cognitive or behavioral correlates that discriminate between good and poor readers. The assumption in both cases seems to be that those factors that discriminate between the good and the bad will serve as likely candidates for subsequent experimental research, research in which those variables are systematically manipulated to determine whether or not improvement occurs.

The research conducted under the auspices of the California Beginning Teacher Evaluation Study (Fisher, Berliner, Filby, Marliave, Cotten, Dishaw, & Moore, 1978), by Brophy and Evertson (1976), and reviewed by

Rosenshine (1979, 1980), and Rosenshine and Berliner (1978) all represent variations on this paradigmatic theme. Also, the debate centering on the follow-through reports (e.g., Becker, 1977; House, Glass, McLean, & Walker, 1978) provides some provocative data regarding effective aspects of instruction. Finally, the work of Stallings (e.g., Stallings, Needies, & Staybrook, 1979) is relevant to this set of issues.

Summarizing almost simultaneously (and hence oversimplifying the situation), we get the following scenario. First, the greater the proportion of time students spend on a task, the better their performance on the task. Academic engaged time, to use Rosenshine and Berliner's (1978) term, is a reasonable predictor of reading achievement gain, ranging in magnitude from correlations of .30 to .59 (e.g., Fisher, et al., 1978; Samuels & Turnure, 1974; Stallings & Kaskowitz, 1974).

A separate variable, related to engaged time, that could be labeled "content covered" or "content measured" (the two tend to be confounded), tends to be positively related to achievement and/or achievement gain (Anderson, Evertson, & Brophy, 1979; Barr, 1973-74; Brown, 1978; Good, Grouws, & Beckerman, 1979; Harris & Serwer, 1968). This relationship seems to hold, across a wide range of content: number of books read, number of words taught, number of basal levels completed, or number of computerized modules mastered.

Third, error rate seems to add a significant amount of power in predicting achievement above and beyond engagement and content covered. The California Beginning Teacher Evaluation Study (Fisher et al., 1978) examined the additional predictive power of error rate over simple engagement and time allocated for reading. They found that error rate increased

the correlation with reading achievement in 7 of 10 predictions. Interestingly, the data suggest that lower error rates (about 80-90% correct) are successful with low achievers whereas somewhat higher error rates (about 70% correct) are more effective with high achievers. These data derive from a variety of settings for teacher-student interactions (words correct, answers to questions correct, etc.). Also, the combined predictions (combining allocated time, engagement rate and error rate) suggest that time spent on decoding is correlated with achievement in Grade 2 to a greater degree than is time spent on comprehension, while the reverse is true in Grade 5. Of course, this may reflect little more than the differences in criterion test items across grades.

Fourth, group instruction, particularly small group instruction, is consistently associated with positive gains in achievement, while individualized instruction is associated with negative or negligible gains (Fisher et al., 1978; Kean, Summers, Ranietz, & Farber, 1979; Soar, 1973; Stallings & Kaskowitz, 1974). Granted, neither is as effective as one-to-one instruction (Smith & Glass, 1980); however, assuming a normal student-teacher ratio (15:1 to 30:1), group instruction appears more effective than individualized seatwork-oriented instruction. Note, however, that grouping is confounded with engagement, which may be the operative variable; for example, in the BTES study (Fisher et al., 1978), engagement rates averaged 84% in group situations and about 70% in individualized situations. Even more dramatic is the data for conscious nonattendance to task: 16% when students worked alone versus 5% when students worked in groups.

All these findings taken together, what emerges is the conclusion that traditional instruction consistently wins out over innovative

instruction. One is tempted to conjure up a picture of a hardhearted taskmaster of a teacher drilling students mercilessly on boring skills, using choral recitation as a major response mode. Such is not the case. Studies that have examined qualitative and affective variables in successful and unsuccessful classrooms tend to have difficulty discriminating between classrooms on these sorts of variables. In fact, most studies have found very little in the way of direct student criticism or harshness to students (e.g., Anderson, Evertson & Brophy, 1979), and such teacher behavior either correlated negatively with achievement (Soar, 1973; Soloman & Kendall, 1976; Stallings et al., 1979) or was positively related to achievement only when the criticism specified desirable alternative behaviors. Remember that these same studies found positive relationships between the four previously reviewed variables and achievement; thus it must be the case that this traditional cluster of teacher strategies does not lead to cold or harsh teacher/student interactions.

I cannot leave this realm of research without commenting on the research evaluating direct instructional models. Most of the debate about the efficacy of direct versus incidental instruction has centered on Follow Through comparisons between DISTAR and other more humanistically oriented programs (House et al., 1978; Stallings & Kaskowitz, 1974; Becker & Carnine, Note 1). Hence the research has been confounded in the sense that it has compared direct instruction in decoding using a fairly regimented group-oriented program (DISTAR) with incidental instruction emphasizing comprehension in a humanistically oriented program (for example, the language experience programs in Follow Through). In short, the direct instructional model (after Engelman, Note 2), which emphasizes rules,

minimal contrasts to elicit discrimination of distinctive features of the rules, and lots of practice after instruction, has not been fairly tested. We do not know from the Follow Through research which aspect of the DISTAR program--direct instruction, decoding, or regimentation--leads to superior decoding performance in high-risk populations.

To foreshadow a later section of this paper, let me mention that the systematic application of direct instructional approaches in the area of comprehension instruction has lead to superior comprehension performance in several studies (Day, 1980; Gordon, 1980; Hansen, 1981; Raphael, 1980; Tharp, Note 3). The Tharp (Note 3) study is important because it has evaluated a frontal assault on comprehension instruction over a several-year period. What is remarkable about the results of Tharp's research is that groups of high-risk native Hawaiian children have moved from mean comprehension test scores hovering near the 20th percentile to near the 60th percentile.

Rosenshine and Stevens (in press) characterize this gestalt of variables as an overall academic orientation to teaching and learning. Success seems to be characteristic of warm but task-oriented classroom environments where students are expected to and do complete work related to reading and reading skill development. The teachers working in these classrooms might well be labeled Hard-nosed Humanists.

Current Practices in Teaching Reading Comprehension

The research surveying current practices for teaching reading comprehension is limited to a single study (Durkin, 1978-79). Durkin observed 17,997 minutes of instruction in both reading and social studies classes. She developed a scheme for classifying teacher behaviors. Comprehension

instruction was limited to activities in which the teachers conducted lessons in which they discussed/interacted with students about how one goes about doing comprehension tasks--finding main ideas, paraphrasizing, determining sequence, etc. Comprehension assessment was represented by teachers quizzing students about stories they had read (and focusing on right answers). Comprehension assignment consisted of mentioning to students how they were to go about completing a workbook, ditto, or other written assignment. There were many other categories, but these are most relevant for our purposes.

Of the total 17,997 minutes of observation, Durkin found that less than 1% was devoted to activities that met one of her definitions of instruction. What were teachers doing in the classes she observed? First, they were giving many assignments for students to do on their own without teacher supervision. Second, they were asking students many questions about stories they read and were focusing on getting THE right answer. Third, they answered a fair number of individual questions about assignments. What was going on in the name of comprehension instruction? Put simply, assignment giving and question asking. The prevailing wisdom concerning comprehension instruction seems to be that if students get enough exposure to a skill or kind of question, they will eventually improve at it. While such a position may be consistent with the engaged-time-on-task argument derived from Rosenshine and Steven's (in press) review, it is not consistent with arguments emanating from the direct instruction or grouping findings. Furthermore, simply on common sense grounds there is something suspicious about a position whose implicit rationale is that if children have trouble with X, what they need is to practice X more often.

Such a position probably works fine for students who can perform the task at a moderate error rate; however, for students who hover near chance level on the task, the additional practice may only reinforce their already misguided strategies. In other words, what Durkin found in our schools in the name of comprehension instruction may be a practice that promotes a "the rich get richer and the poor get poorer" syndrome.

In a sequel to her classroom observation study, Durkin (1981) examined the teachers' editions of five currently popular basal reading programs, looking for instances of comprehension instruction defined in terms comparable to the criteria used in her earlier study (Durkin, 1978-79). While the sheer incidence of comprehension instruction was higher than in her previous study, the general pattern of a dominant reliance on assessment and mentioning was replicated.

Durkin's two studies, taken together, reveal a picture of virtually no direct instruction in comprehension. Instead, teachers seem to spend most of their classroom discussion time asking students questions about stories they have read and giving assignments. Regarding comprehension skills--such as main idea, sequence, cause-effect, fact-opinion--manuals provide little guidance concerning how the skills ought to be presented to students; teachers apparently provide little guidance to students about how they ought to solve problems and/or answer questions exemplifying these skills. The prevailing wisdom is to provide massive doses of unguided practice. Nor is there much evidence, either in manuals or classrooms, that much goes on in the name of substantive feedback that would allow students to evaluate how well they were performing a task or, more important, what inappropriate strategies they might be adopting. The student

who is not doing well on a particular comprehension skill seems to have little help to look forward to, save additional opportunities to improve performance on his or her own through practice.

Research on Comprehension Instruction

Durkin's two studies tend to engender an atmosphere of pessimism. Perhaps they should. They have probably provided the reading profession a definite service, for they prompt the question, What is the alternative to practice and assessment? As an antidote to that pessimism, let me turn to a review of a few recent studies that have evaluated the effects of direct explicit attempts to help students develop heuristic strategies (if not rules) for dealing with a range of comprehension tasks typically required in schools.

These studies share a set of features. First, all of them are derived directly from basic research on the reading process; that is, they represent attempts to bridge the gap from basic research to a real instructional issue. Second, all have evaluated the efficacy of their instructional treatments by using transfer tasks; they have asked the question, What happens to student performance when instructional crutches are removed? Third, all have obtained positive results; they have shown that the intervention at issue elicits positive gains in some aspect of comprehension. Fourth, all have attended, at least in some way, to the question of control processes. They have included, directly or by implication, techniques that allow students to monitor for themselves whether or not they understand task demands or know when they are performing the task appropriately.

In the first study, Hansen (1981) was interested in ameliorating children's ability and predisposition to draw inferences. Beginning with

the observation that children were best at answering the kinds of questions teachers ask most often, i.e., literal recall of story details, she wondered whether this observation represented a robust developmental trend, an accident of children's instructional history (i.e., they have more practice at literal questions), or a fact about the world (literal questions are inherently easier than inferential questions).

She devised three instructional treatments. In the first, a business-as-usual approach, average second-grade students were given a traditional diet of questions accompanying their basal reader stories--about 80% literal to 20% inferential questions. In the second, a practice-only treatment, literal questions were removed from these children's basal reader lives altogether; they received only inferential questions. In the third, students received the traditional question diet but were confronted lesson after lesson with pre-reading strategy designed to help them process new (text) information in light of existing (head) knowledge structures. Prior to each story, they were asked to predict what they would do and what the story protagonist would do when either confronted 2 or 3 critical situations (actual situations from the story to be read). They then read the story to compare their predictions with what actually occurred (a la Directed Reading-Thinking Activity). In addition, they were provided with a visual model of comprehension as a process of relating the new to the known.

Four kinds of dependent measures were analyzed, using pretest story understanding tasks (answering literal and inferential probes) as a covariate in a multivariate ANOVACOVA. On the first measure, literal and inferential probes from the last five stories in which the instruction was

embedded, both the practice-only and the strategy training group outperformed the traditional group on both literal and inferential probes. In addition, where differences existed between the two experimental groups, they favored the strategy training group. The data suggest that a set for inferential processing induces a levels-of-processing effect that generalizes to both inference and literal tasks, at least in the local environment of the stories in which the instruction was embedded.

On the second measure, literal and inferential probes from totally new and unaided stories, the two inference-oriented groups exceeded the traditional group only on inference probes for the familiar transfer story. These data suggest that whatever heuristic developed could not overcome the strong influence that prior knowledge has on inference performance (i.e., no differences on the inference probes for the topically unfamiliar selection).

On the third measure, free recall of a totally new story, there were absolutely no differences, arguing for a transfer-of-identical-elements phenomenon. In short, since the students never practiced free recall, their ratio of intrusions (inferences) to text reproductions was not influenced.

On the fourth measure, a posttest only standardized reading test, there was a treatment by subtest interaction. On the vocabulary subtest, there were no reliable differences among groups, strengthening the argument that there were no pre- or postexperimental general verbal ability differences among the groups. On the comprehension subtest, however, there were strong differences favoring both experimental groups over the traditional group. At first blush this may seem surprising, since standardized tests

are typically insensitive to specific instructional treatments. However, the standardized test used was the Stanford Achievement Test, which uses a modified cloze (fill-in-the-blank) response format. Such a format, if it does anything, places a premium on inferences to prior knowledge; how else would anyone determine the best fit for the cloze blank. Hence, the transfer is not so surprising.

The primary conclusion one can draw from these data is that inference ability, even for young students, is amenable to direct training and monitoring; however, the local and task-alike transfer effects are more impressive than the broad transfer effects.

Gordon (1980) extended, at least in part, the inference training hypothesis to older children (Grade 4). Over a period of eight weeks, she contrasted the effects of an even more explicitly trained inference group with a placebo control group that received fun language experience and immersion activities and a second experimental group whose instruction focused on activating and fine-tuning preexisting content schemata (the topics addressed in the stories) and structure schemata (helping students develop an abstract framework for what is entailed in a story) before and after reading.

Five dependent measures were used: (a) comprehension of literal and inferential probes summed over the eight stories in which the instruction was embedded, (b) comprehension of literal and inferential probes on transfer stories read immediately following the eight-week experiment, (c) same as (b) but delayed two weeks, (d) a standardized comprehension test measure, and (e) free recall protocols from the last story read in the training period.

While the results are not quite so dramatic as in the Hansen (1981) study, the patterns of significant results are consistent. There were no significant differences between groups on the standardized test or on the immediate comprehension test, again suggesting that broad transfer is difficult to obtain. However, there were statistically reliable differences favoring the inference training group on inference items derived from the instructional stories. Also, high-achieving but not low-achieving students in that group did better than other groups on the inference items on the delayed posttest. The most remarkable differences favored the content and structure schemata activation group on the free recall protocols; their scores were often two or three standard deviations above the inference group and the placebo control group, particularly on recall measures which were sensitive to the development and use of a story schema. Apparently these students developed an abstract story "map" which served them well in encoding and retrieving information structurally important in a story schema. As with the Hansen and Pearson study, one is more impressed with the local than the broad transfer effects. Also, one is struck by the specificity of the transfer that does occur; the principle of transfer of identical elements (the greater the similarity between training and transfer tasks, the greater the likelihood of transfer) suggests itself. One is tempted also to invoke Rosenshine's engaged-time-on-task principle in explaining these data.

Hansen and Pearson (in press) have followed up earlier inference training research with modified techniques and different populations. In earlier research, (Hansen, 1981) contrasted a strategy approach with a practice-only approach and a business-as-usual control condition. In the

follow-up, Hansen and Pearson combined strategy and practice into a single treatment to be contrasted with the conventional approach. They also trained four teachers to administer the treatments instead of teaching the classes themselves, as had been done earlier. Finally, they used good and poor fourth-grade readers instead of average second-grade students.

The combined approach proved not to be advantageous for good readers in comparison to the control group; however, it proved remarkably effective for the poor readers. Experimental poor readers exceeded their control counterparts on inference measures taken from the materials in which the instruction was embedded as well on measures from three transfer passages for which no instruction was offered. In fact, when all students read and answered questions from a common transfer passage, poor experimental students reading at a 3.1 level scored as well as good control students reading at a 6.2 level. From these data, and the data from the earlier study, they concluded that younger and poorer readers benefit from conscious explicit attempts to alter comprehension strategies; older good readers, on the other hand, seem not to benefit, perhaps because they are capable of developing adequate strategies on their own.

Raphael (1980) cast the inference training paradigm directly in a more general approach to question-answering. Over four 45-minute sessions she trained average 4th-, 6th-, and 8th-grade students (also low, average, and high 6th-grade students) to monitor their allocation of resources (information in the text versus knowledge stored in memory) in generating answers to questions that invited textually explicit comprehension (deriving an answer from the same text sentence from which the question was generated),

textually implicit comprehension (deriving an answer from a text sentence different from the one from which the question was derived), or scriptally implicit comprehension (deriving an answer from one's store of prior knowledge). She modified this scheme, taken from Pearson and Johnson (1978), for students by labeling the three response types RIGHT THERE, THINK AND SEARCH, and ON MY OWN, respectively.

Using a model —> guided practice —> independent practice —> direct feedback instructional design, she guided the students to apply the strategy to increasingly larger text segments (one paragraph to a 600-word passage) with an increasingly larger number of questions per lesson and increasingly fewer feedback prompts from the instructor. In the strategy, students read the relevant text and the question, generated an answer, and then decided which of the three strategies they had used to generate the answer.

In the transfer test, students read entirely new passages on their own, answered questions, and decided on the strategy they thought they had used to generate the answer. The performance of the training group was contrasted not with an untreated control but with a control group that received a 20-minute orientation to the response classification task. Four dependent measures were analyzed: (a) hits (Did the student give his response strategy the same category rating as the experimenter thought was the most readily invited strategy given the particular question and text--in other words, did the student judge himself to do what the experimenter thought most students would do?), (b) matches (irrespective of response quality, did the student actually do what she said she did?), (c) appropriate responses (Did the student give a response that, either because of direct selection from the target position or through a chain of logical and/or pragmatic

reasoning from the target proposition, could be scored correct given a complex set of scoring protocols that allowed for considerable deviation from the expected response, and (d) correct hit matches (given that student achieved a hit [Did what the experimenter expected] and a match [Did what she said she did] what was the probability that she got the item correct?).

On all of these response measures, reliable differences were found favoring the training group over the orientation group; that is, trained students got better at discriminating task demands of different types of questions, evaluating their own behavior, and giving quality responses. Moreover on the conditional measure, which requires discrimination and evaluation and response quality, training/orientation differences were magnified even further. Apparently students changed both their response strategies and their response monitoring strategies. Raphael concluded that they had developed both new comprehension and comprehension monitoring strategies that gave them more control over a traditional but pervasive question answering task.

Working with low-ability community college students, Day (1980) contrasted approaches to training students to write summaries for prose passages. The treatments differed systematically from one another in terms of how rules for writing summaries were integrated with self-management strategies designed to help students monitor their own progress in summary writing. Treatment 1 consisted of self-management alone (a fairly traditional self-checking procedure to determine whether the summary conveyed the information the student intended to convey). Treatment 2 was rules alone; that is, subjects were trained to use van Dijk and Kintsch's (1978)

five rules for summarizing narratives: delete redundancy, delete irrelevancies, subordinate subtopics, select topic sentences, create topic sentences. Treatment 3 simply put Treatments 1 and 2 together in sequence. First do one; then, the other. Treatment 4 integrated the rules and self-management strategies into a single coherent routine. One might say that the four treatments varied along a continuum of integration of explicit training and explicit monitoring devices. A model —> feedback —> practice instructional design was used.

The dependent measure was the proportion of time students used each of the five summarization rules (number of actual uses/number of potential opportunities to use). Day found that from pretest to posttest there was a ceiling effect on the two deletion rules; that is, almost all students could already apply them. On the subordination rule, all but Treatment 1 (self-management alone) students made significant gains, with the greatest gains accruing to the integrated group (Treatment 4). On the selection rule again Treatments 2, 3, and 4 exhibited greater gain than did Treatment 1; however, there were no reliable differences among Treatments 2-4. Also, average-ability students gained more than low-ability students. On the creation rule, a pattern similar to that found for subordination emerged: The greatest gains accrued to the integrated group (Treatment 4). Furthermore, posttest performance indicated that while pre-post gains were similar across rules, absolute performance levels were conditioned by rule complexity: Rule 3 > Rule 4 > Rule 5.

Day's data suggest that with different tasks and with slower students, ". . . explicit training in strategies for accomplishing a task coupled

with routines to oversee the successful application of those strategies is clearly the best approach"(p. 15).

This summary provided by Day could well serve as a summary for the four studies reviewed in this section. All point to the direction of making clear what the task requirements are, providing heuristic guidelines for task completion, allowing substantial massed practice along with substantive feedback, and insuring some provision for self-monitoring. The data are encouraging. It looks as though we can teach comprehension skills after all.

The last section clearly reveals my own biases about the direction instructional research on reading comprehension ought to take. Research should focus on explicit attempts to help students develop independent strategies for coping with the kinds of comprehension problems they are asked to solve in their lives in schools. It is interesting that one could probably infer that such research was needed by examining the gaps in instruction found by Durkin (1978-79) and the positive correlations between existing instructional practices and achievement noted by people like Rosenshine and Stevens (in press). That the few instructional studies on reading comprehension also support such a line of research is encouraging.

As a general model for how we might proceed, let me offer a set of guidelines paraphrased from Brown, Campione, and Day (1981):

1. The trained skill must be instructionally relevant.
2. Training should proceed from simple to complex.
3. An analysis of training and transfer tasks should provide evidence of where breakdowns occur.

4. There should be explicit instruction concerning when and how to use the strategies.
5. Feedback should be given during class discussions and for independent work.
6. A variety of passages (or other materials) should be used in order to facilitate transfer to new situations.
7. Self-checking procedures should be used as an inherent part of operationalizing the training strategy.

In reflecting upon these guidelines and the studies which show the value of direct explicit attempts to improve comprehension skills, I am struck by the consistency of this perspective with what we might call common sense. The questions of interest then become How did we lose our common sense? and How do we find it once more?

In speculating upon the loss, I am convinced that instruction somehow got lost within the prevailing emphasis upon sophisticated materials and management schemes in the decade of the 70's. Never before have we had such an array of texts, workbooks, worksheets, games, and kits available to teach reading skills. In such a mileau it may be seductive for educators to believe that materials really do teach. In fact, a recent survey by Shannon (1981) confirms such a belief among teachers and (even more strongly) administrators.

In anticipating a return to what we might want to label the science and art of teaching (as opposed to managing), I think the justification exists for placing more emphasis on direct explicit teaching, interactive discussions, substantive feedback, and control and self-monitoring strategies. Hopefully, as we accumulate additional evidence supporting the efficacy

of these techniques, particularly in natural classroom environments, and as we discuss these techniques with practitioners, we will return to the model of teacher as teacher, and perhaps the return will occur as naturally and quickly as we turned to the model of teacher as manager during the 1970's.

Reference Notes

1. Becker, W. C., & Carnine, D. Direct instruction--A behaviorally based model for comprehensive educational intervention with the disadvantaged. Unpublished manuscript, University of Oregon, 1978.

2. Engleman, S. Sequencing cognitive and academic tasks. Unpublished paper, University of Oregon, 1977.

3. Tharp, R. The direct instruction of comprehension: Results and description of the Kamehameha Early Education Program. Paper presented at the annual meeting of the American Educational Research Association. Boston, April 1980.

References

Anderson, R. C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. Frameworks for comprehending discourse. American Educational Research Journal, 1977, 14, 367-382.

Anderson, R., Evertson, C., & Brophy, J. An experimental study of effective teaching in first grade reading groups. Elementary School Journal, March 1979, 193-222.

Baker, L., & Brown, A. Comprehension monitoring and critical reading. In J. Flood (Ed.), Comprehension and cognitive processes. Newark, Del.: International Reading Association, in press.

Barr, R. Instructional pace differences and their effect on reading acquisition. Reading Research Quarterly, 1973-74, 9, 526-554.

Becker, W. C. Teaching reading and language to the disadvantaged--What we have learned from field research. Harvard Educational Review, 1977, 47, 518-543.

Brown, A. L. Knowing when, where, and how to remember: A problem of metacognition. In R. Glasser (Ed.), Advances in instructional psychology. Hillsdale, N.J.: Erlbaum, 1978.

Brown, A. L., Campione, J. C., & Day, J. Learning to learn: On training students to learn from texts. Educational Researcher, 1981, 10, 14-24.

Brophy, J., & Evertson, C. Learning from teaching: A developmental perspective. Boston: Allyn & Bacon, 1976.

Chi, M. T. H. Knowledge structures and development. In R. S. Siegler (Ed.), Children's thinking: What develops? Hillsdale, N.J.: Erlbaum, 1978.

Day, J. D. Training summarization skills: A comparison of teaching methods. Unpublished doctoral dissertation, University of Illinois, 1980.

Durkin, D. What classroom observations reveal about reading comprehension instruction. Reading Research Quarterly, 1978-79, 14, 481-533.

Durkin, D. Reading comprehension instruction in five basal reading series. Reading Research Quarterly, 1981, 16, 515-544.

Dijk, T. A. van, & Kintsch, W. Cognitive psychology and discourse. In W. U. Dressler (Ed.), Trends in text linguistics. New York: DeGruyter, 1978.

Fisher, C., Berliner, D., Filby, N., Marliave, R., Cohen, L., Dishaw, M., & Moore. Teaching behaviors, academic learning time and student achievement: An overview. In C. Denham & A. Lieberman (Eds.), Time to learn. Washington, D.C.: National Institute of Education, May 1978.

Good, T., Grouws, D. A., & Beckerman, T. M. Curriculum pacing: Some empirical data in mathematics. Journal of Curriculum Studies, 1979, 11, 263-281.

Gordon, C. J. The effects of instruction in metacomprehension and inferencing on children's comprehension abilities. Unpublished doctoral dissertation, University of Minnesota, 1980.

Hansen, J. The effects of inference training and practice on young children's comprehension. Reading Research Quarterly, 1981, 16, 391-417.

Hansen, J., & Pearson, P. D. An instructional study: Improving the inferential comprehension of good and poor fourth-grade readers (Tech. Rep.). Urbana: University of Illinois, Center for the Study of Reading, in press.

Harris, A. J., & Serwer, B. L. The craft project: Instructional time. Reading Research Journal, 1968, 2(1), 27-68.

House, E. R., Glass, G. V., McLean, L. D., & Walker, D. F. No simple answer: Critique of the follow through evaluation. Harvard Educational Review, 1978, 48, 128-160.

Kean, M., Summers, A., Ranietz, M., & Farber, I. What works in reading. Philadelphia: School District of Philadelphia, 1979.

Mandler, J. M. A code in the node: The use of a story schema in retrieval. Discourse Processes, 1978, 1, 14-35.

Meyer, B. J. F. What is remembered from prose: A function of passage structure. In R. O. Freedle (Ed.), Discourse production and comprehension (Vol. 1). Norwood, N.J.: Ablex, 1977.

Meyer, B. J. F., Brandt, D. M., & Bluth, G. J. Use of top level structure in text: Key for reading comprehension. Reading Research Quarterly, 1980, 16, 72-103.

Neilsen, A. The role of macrostructures and relational markers in comprehending familiar and unfamiliar written discourse. Unpublished doctoral dissertation, University of Minnesota, 1977.

Omanson, R. C. An analysis of narratives: Identifying central, supportive, and distracting content. Discourse Processes, in press.

Paris, S. C., & Lindauer, B. K. The role of inference in children's comprehension and memory. Cognitive Psychology, 1976, 8, 217-227.

Paris, S. C., & Upton, L. R. Children's memory for inferential relationships in prose. Child Development, 1976, 47, 660-668.

Pearson, P. D., Hansen, J., & Gordon, C. The effect of background knowledge on young children's comprehension of explicit and implicit information. Journal of Reading Behavior, 1979, 11, 201-210.

Pearson, P. D., & Johnson, D. D. Teaching reading comprehension. New York: Holt, Rinehart, & Winston, 1978.

Pichert, J. W. Sensitivity to what is important in prose (Tech. Rep. No. 149). Urbana: University of Illinois, Center for the Study of Reading, November 1979. (ERIC Document Reproduction Service No. ED 179 946)

Pichert, J. W., & Anderson, R. C. Taking different perspectives on a story. Journal of Educational Psychology, 1977, 69, 309-315.

Raphael, T. E. The effects of metacognitive strategy awareness training on students' question answering behavior. Unpublished doctoral dissertation, University of Illinois, 1980.

Rosenshine, B. Content, time, and direct instruction. In H. Walberg & P. Peterson (Eds.), Research on teaching: Concepts, findings, and implications. Berkeley, Calif.: McCutchan, 1979.

Rosenshine, B. How time is spent in elementary classrooms. In C. Denhan & A. Lieberman (Eds.), Time to learn. Washington, D.C.: National Institute of Education, May 1980.

Rosenshine, B., & Berliner, D. Academic engaged time. British Journal of Teacher Education, 1978, 4, 3-16.

Rosenshine, B. V., & Stevens, R. Advances in teacher education research. Journal of Special Education, in press.

Rumelhart, D. E. Towards an interactive model of reading. In S. Dornič (Ed.), Attention and performance VI. Hillsdale, N.J.: Erlbaum, 1977.

Samuels, S. J., & Turnure, J. E. Attention and reading achievement in first grade boys and girls. Journal of Educational Psychology, 1974, 66, 29-32.

Shannon, P. Teachers' self-perceptions and reification of instruction within reading instruction. Unpublished doctoral dissertation, University of Minnesota, 1981.

Smith, M., & Glass, G. Meta-analysis of research on class size and its relation to attitudes and instruction. American Educational Research Journal, 1980, 17, 419-434.

Soar, R. S. Follow-through classroom process measurement and pupil growth (1970-71): Final report. Gainesville: University of Florida, College of Education, 1973.

Soloman, D., & Kendall, H. J. Final report: Individual characteristics and children's performance in varied educational settings. Chicago: Spencer Foundation Project, May 1976.

Stallings, J. A., & Kaskowitz, D. Follow-through classroom observation evaluation 1972-73. Menlo Park, Calif.: Stanford Research Institute, 1974.

Stallings, J. A., Needles, M., & Staybrook, N. How to change the process of teaching basic reading skills in secondary schools. Menio Park, Calif.: Stanford Research Institute International, 1979.

Stanovich, K. D. Toward an interactive-compensatory model of individual differences in the development of reading fluency. Reading Research Quarterly, 1980, 16, 32-71.

Stein, N., & Glenn, C. G. An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), New directions in discourse processing. Norwood, N.J.: Ablex, 1979.

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

Trabasso, T. On the making and assessment of inferences during reading. In J. T. Guthrie (Ed.), Reading comprehension and education. Newark, Del.: International Reading Association, 1981.