A Quasi-Experimental Validation of Transactional Strategies Instruction With Low-Achieving Second-Grade Readers

Rachel Brown
State University of New York at Buffalo

Michael Pressley
University at Albany, State University of New York

Peggy Van Meter
University of Maryland—College Park

Ted Schuder
Maryland State Department of Education

Second-grade, low-achieving students experienced a year of either transactional strategies instruction or highly regarded, more conventional second-grade reading instruction. By the end of the academic year, there was clear evidence of greater strategy awareness and strategy use, greater acquisition of information from material read in reading group, and superior performance on standardized reading tests by the transactional strategies instruction students. This is the clearest validation to date of educator-developed transactional strategies instruction.

Since Durkin's (1978–1979) seminal discovery that American students received little instruction about how to comprehend text, there have been extensive efforts to identify strategies that can be taught to students to increase their understanding of and memory for text. Early strategy research (for reviews, see Dole, Duffy, Roehler, & Pearson, 1991; Pressley, Johnson, Symons, McGoldrick, & Kurita, 1989) tended to focus on instruction of individual strategies and improvements in narrowly defined performances (e.g., improvement on standardized comprehension tests when reading strategies were taught). The typical research tactic taken in these studies was to teach one group of students to use a particular cognitive strategy while reading, often a strategy consistent with a theory of knowledge representation favored by the researcher, with control students left to their own devices to understand text as best they could.

Through this approach, a relatively small number of individual strategies were proved effective in increasing elementary students' comprehension of and memory for text (e.g., visualizing ideas in text, summarizing, and self-questioning). What the single-strategy investigations demonstrated was that if students were under exceptionally strong instructional control (i.e., they were told to carry out a particular strategy on a particular occasion), they could carry out strategies that would improve comprehension and learning. Seldom was generalized use of individual instructed strategies observed, nor was there evidence of generalized improvement in reading.

On the basis of what is now known about skilled reading, it is not surprising that improvement in reading has required more than instruction in single strategies. During the late 1970s and early 1980s, a number of analyses of skilled reading were conducted (e.g., Johnston & Afflerbach, 1985; Lytle, 1982; Olshavsky, 1976–1977; Olson, Mack, & Duffy, 1981; see Pressley & Afflerbach, 1995, for a summary). What became apparent was that skilled reading did not involve the use of a single potent strategy but rather orchestration of cognitive processes. This understanding—that skilled readers coordinate a number of strategies while reading—partially fueled researcher efforts to develop instructional interventions that involved teaching of multiple comprehension strategies (Baker & Brown, 1984).

A well-known researcher-designed, multiple-strategies instructional package was Palincsar and Brown's (e.g., 1984) reciprocal teaching. Palincsar and Brown taught students to apply four strategies to expository text as they read (generate predictions, ask questions, seek clarification, and summarize content). The students used these strategies in reading groups, with the adult teacher releasing control of the strategic processing as much as possible to the group. Palincsar and Brown's prediction, consistent with Vygotsky's (e.g., 1978) theory of socially mediated learning, was that participation in reading group discussions that involved predicting, questioning, seeking clarification, and

summarizing would lead to internalization of these processes by group members. In fact, a month or two of such instruction produces noticeable improvement in the use of the focal strategies but only modest improvement on standardized reading tests (for a review, see Rosenshine & Meister, 1994).

In addition to the research of Palincsar and Brown (1984), there were other attempts to teach multiple comprehension strategies. Some involved presenting a large number of strategies quickly; these approaches typically failed to produce improvements in elementary-level readers' comprehension (e.g., Paris & Oka, 1986). Other interventions involved more intensive direct explanation and modeling of small repertoires of strategies; these approaches generally were more successful in improving reading (e.g., Bereiter & Bird, 1985; Collins, 1991; Duffy et al., 1987).

Many educators became aware of strategy researchers' instructional successes and began to import strategies instruction into classrooms. What became apparent, however, was that when strategies instruction was successfully deployed in schools, it involved much more than the operations studied in the well-controlled experiments (Pressley, Goodchild, Fleet, Zajchowski, & Evans, 1989). This factor motivated Pressley and his colleagues to study extensively how elementary educators implemented comprehension strategies instruction in schools (see Pressley & El-Dinary, 1993).

After investigating several educator-developed programs, our research group proposed that effective elementary-level comprehension was "transactional" in three senses of the term (Pressley, El-Dinary, et al., 1992). First, readers are encouraged to construct meaning by using strategies that enable the linking of text content to prior knowledge, consistent with Rosenblatt's (1978) use of the term. Second, much of the strategies instruction occurs in reading groups, with group members using strategies to construct meaning together. As such, meaning-making is transactional in the sense that the constructed group understanding differs from the personalized interpretations individuals would have generated on their own, especially if they did not use strategies. This is consistent with the use of the term in organizational psychology (e.g., Hutchins, 1991). Third, the teacher's or group members' actions and reactions cannot be anticipated when the reading group uses strategies to construct interpretations. Rather, the responses of all members of the group (including the teacher) are determined in part by those of others in the group, which is a transactional situation according to social development researchers such as Bell (1968). Thus, group members co-determine each other's thinking about text. Because the strategy instruction the research group observed was so "transactional" in these three senses of the term, this type of instruction was called *transactional strategies instruction* (TSI; Pressley, El-Dinary, et al., 1992).

The short-term goal of TSI is the joint construction of reasonable interpretations by group members as they apply strategies to texts. The long-term goal is the internalization and consistently adaptive use of strategic processing whenever students encounter demanding text. Both goals are

promoted by teaching reading group members to construct text meaning by emulating expert readers' use of comprehension strategies: to emulate how expert readers constructively respond when they need to understand challenging text (e.g., Pressley & Afflerbach, 1995; Wyatt et al., 1993). Expert readers are planful and goal-oriented when they read, combine their background knowledge with text cues to create meaning, apply a variety of strategies (e.g., from seeking the important information in text to noting details), monitor their comprehension, attempt to solve their comprehension problems, and evaluate their understanding and performance (e.g., Is the content believable? Is the piece well written? Am I achieving my goals?). The result is a personalized, interpretive understanding of text.

A variety of qualitative methods were used in the previous studies of transactional strategies instruction (see Pressley, El-Dinary, et al., 1992). These included (a) ethnographies; (b) interviews involving questions emanating from Pressley, Goodchild, et al.'s (1989) tentative description of strategies instruction; (c) interviews constructed to illuminate observations made in program classrooms; (d) long-term case studies; and (e) analyses of classroom discourse. Although the TSI programs differed in their particulars, there were a number of common components (Pressley, El-Dinary, et al., 1992):

- Strategy instruction is long-term, with effective strategies instructors offering it in their classroom throughout the school year; the ideal is for high quality process instruction to occur across school years.
- Teachers explain and model effective comprehension strategies. Typically, a few, powerful strategies are emphasized.
- The teachers coach students to use strategies on an as-needed basis, providing hints to students about potential strategic choices they might make. There are many mini-lessons about when it is appropriate to use particular strategies.
- Both teachers and students model use of strategies for one another, thinking aloud as they read.
- Throughout instruction, the usefulness of strategies is emphasized, with students reminded frequently about the comprehension gains that accompany strategy use. Information about when and where various strategies can be applied is commonly discussed. Teachers consistently model flexible use of strategies; students explain to one another how they use strategies to process text.
- The strategies are used as a vehicle for coordinating dialogue about text. Thus, a great deal of discussion of text content occurs as teachers interact with students, reacting to students' use of strategies and prompting additional strategic processing (see especially Gaskins, Anderson, Pressley, Cunicelli, & Satlow, 1993). In particular, when students relate text to their prior knowledge, construct summaries of text meaning, visualize relations covered in a text, and predict what might transpire in a story, they engage in personal interpretation of text, with these personal

interpretations varying from child to child and from reading group to reading group (Brown & Coy-Ogan, 1993).

Although the qualitative studies provided in-depth understanding of the nature of transactional strategies instruction programs, and a variety of informal data attested to the strengths of these programs (e.g., correlational, nonexperimental, and quasi-experimental comparisons conducted by school-district officials; see Brown & Pressley, 1994), what was lacking until the study reported here was conducted were formal comparisons on a variety of reading measures of students who received transactional strategies instruction versus more conventional instruction. This study begins to fill that gap. There were several important challenges to making such comparisons, however. One challenge was determining what should be measured. Reading strategies instruction has tended to focus on gains on one or a few traditional measures of reading performance (Pressley, El-Dinary, et al., 1992). It became clear on the basis of the qualitative studies that transactional strategies instruction probably affected student cognition in a number of ways, however, with both short-term and long-term impacts (Pressley, Schuder, Teachers in the Students Achieving Independent Learning Program, Bergman, & El-Dinary, 1992).

A second challenge was that many indicators in the qualitative research conducted on transactional strategies instruction suggested that the effects of such an intervention appeared in the long term; that is, at a minimum, only after a semester to an academic year of such instruction (see Marks et al., 1993; Pressley, El-Dinary, et al., 1992; Pressley, Schuder, et al., 1992). A credible evaluation had to be long term. A constraint was that students often move in and out of schools at a high rate; thus, holding together large groups of students for several years was impractical. Our solution was to evaluate 1 year of transactional strategies instruction, because 1 year of intervention was all we believed could be completed in the participating district with an intact sample of students.

A third challenge resulted in our decision not to assign teachers randomly to conditions. Becoming an effective transactional strategies instruction teacher takes several years (e.g., El-Dinary & Schuder, 1993; Pressley, et al., 1991; Pressley, Schuder, et al., 1992). Thus, we felt we could not take any group of teachers and randomly assign them to transactional strategies instruction or control conditions. Moreover, we decided not to assign accomplished transactional strategies instruction teachers randomly to teach strategies versus some other approach. Because the teachers were committed to strategies instruction, we felt it was inappropriate to ask them to alter their teaching for an entire year. Our solution was to use a quasi-experimental design involving accomplished TSI teachers and other teachers in the same district, teachers with reputations as excellent reading educators whose instruction followed the guidelines of the district's regular literacy curriculum.

Before proceeding with a description of the formal methods in our study, we summarize some of the most important features of the Students Achieving Independent Learning

(SAIL) program, the specific educator-developed transactional strategies instruction approach evaluated here. A description of SAIL will permit readers to understand our expectations in this quasi-experiment.

The SAIL Comprehension Strategies Instructional Program

The purpose of SAIL is the development of independent, self-regulated meaning-making from text. The program was developed over the course of a decade in one mid-Atlantic school district (see Schuder, 1993, for a history of SAIL and its evolution). SAIL students are taught to adjust their reading to their specific purpose and to text characteristics (Is the material interesting? Does it relate to the reader's prior knowledge?). SAIL students are instructed to predict upcoming events, alter expectations as text unfolds, generate questions and interpretations while reading, visualize represented ideas, summarize periodically, and attend selectively to the most important information. Students are taught to think aloud (e.g., Meichenbaum, 1977) as they practice applying comprehension strategies during reading group instruction. For example, they reveal their thinking to others when they talk about their past experiences in relation to text. All of these reading processes are taught as strategies to students through direct explanations provided by teachers, teacher modeling, coaching, and scaffolded practice, both in reading groups and independently.

Direct explanations and modeling of strategic reasoning are critical components for preparing students to internalize and use strategies adaptively. These core components start the long-term process of helping students become more self-regulated and skillful readers. Direct explanations include providing students with information about the benefits of strategy use, as well as when and where to use strategies. In this excerpt, a SAIL teacher explained what is necessary to make good predictions:

T: We're going to set a couple of goals. So let's listen carefully and really really try to meet these goals by the end of the lesson. The first thing I want everybody to try to do is to make really good predictions. Can anybody tell me what a prediction is? S8?

S8: When you think what's gonna happen next.

T: When you try to guess what's going to happen next. If you're going to be a good predictor, how do you make good predictions? What do you need to have? S5?

S5: Enough information.

T: You have to have enough information in order for a good prediction to be made. Where can you find your information to get a prediction, where can you find it, S8?

S8: In the book?

T: In the book? You mean like, from what you've already read?

S8: Yeah.

T: S6?

S6: In your head.

T: In your head. Sometimes, a fancy word for that is background knowledge. In other words, knowledge means that you know. If you already know something about

foxes, or about what a trot is, you might be able to even make a prediction about what the story is about. But maybe we should read a little bit into it to get a little more information to make sure we can make some good predictions.

Modeling, another critical SAIL component, does not consist solely of showing students how to use a strategy. Instead, SAIL teachers verbally explain their thinking and reasoning as they model appropriate use of strategies. In the following example, a SAIL teacher modeled her use of strategies, verbalizing her thinking as she applied a strategy in response to the demands of the text and her need to understand. However, before modeling, she explained to students why she was going to model: to help them observe how and why she used strategies to comprehend what she was reading:

> T: I'm gonna start this morning modeling like I usually do . . . . This is gonna be a real good opportunity for you to use a lot of your strategies. There are a lot of big words in this story. Okay, so it's going to give us a chance to use some of our "fix-up kit" and it's gonna also give us a chance to use a lot of background knowledge, things that we already know from our own life, to help us understand what this story is about . . . ,

The teacher proceeded to read "Fox was a fine dancer. He could waltz, he could boogie, he could do the stomp." She then modeled her thinking for students:

> T: You know what? I'm thinking waltzing, boogieing, doing the stomp. I don't really know what the stomp is. But I'm thinking to myself that the stomp must be a dance because I do know that the waltz is a dance. That's when two people dance together. Because I used to see that on the Lawrence Welk Show. My grandmothers used to watch that a lot. And the boogie, well, I know that was a dance when I was in high school and that's when you move real fast. So, I'm thinking, the word *stomp,* I know that, well you can stomp your foot, and maybe that's what people do when they do the stomp. But I still think it's a dance. So that's what I'm gonna think, that. He could do the stomp, so that's a dance.

The teacher related information from her own experiences to text details. She used her prior knowledge to apply one of the "fix-up" strategies, guessing, when figuring out the meaning of an unknown word (i.e., *stomp*). Later in the same lesson, a student substituted a word when she came to a word she did not know. The teacher reminded students of the strategic reasoning she used when she initially verbalized her thinking. She then explained how readers can select different "fix-up" strategies (i.e., substituting a known word for an unknown one or relying on picture clues) to achieve the same purpose: to understand the gist of a passage.

> T: But, back over here, when I first said that there were some big vocabulary that had to do with the kinds of dances . . . we were talking about the fact that S1 substituted as a strategy and she was able to figure out . . . [what the word meant]. And then S6 over there gave her the real word, and then we found out that the real word wasn't that important because we could understand.

Then I was thinking back to the fact that I didn't know what the stomp was and here I was looking at the picture clue. So, even though S1 was having trouble with the vocabulary, she could still get the gist of the type of dances by looking at the three pictures. There were three types of dances and there were three pictures.

In addition, SAIL students are taught multiple methods for dealing with difficult words, including skipping them, using context clues to determine the meaning of hard-to-decode and unfamiliar words, and rereading for more clues to meaning. Overreliance on any one strategy is discouraged. For example, skipping every unknown word can lead to comprehension failures, particularly if the skipped words are central to the meaning of the text. Instead, students are taught that skipping is just one of several strategies available to them when they encounter unknown words. When students are taught to ignore unknown words judiciously, skipping becomes a powerful problem-solving strategy for those who otherwise might linger too long over an undecodable word. In general, students are taught that getting the overall meaning of text is more important than understanding every word, so that difficult words sometimes can be skipped with little or no loss in meaning.

When SAIL instruction occurs in reading groups, it differs in a number of ways from more conventional reading group instruction: (a) Prereading discussion of vocabulary is eliminated in favor of discussion of vocabulary in the context of reading. (b) The almost universal classroom practice of asking comprehension-check questions as students read in group (e.g., Mehan, 1979) is rarely observed in transactional strategies instruction groups (Gaskins et al., 1993). Instead, a teacher gauges literal comprehension as students think aloud after reading a text segment. (c) There are extended interpretive discussions of text, with these discussions emphasizing student application of strategies to text.

Although reading group is an important SAIL component, the teaching of strategies extends across the school day, during whole-class instruction, and as teachers interact individually with their students. Reading instruction is also an across-the-curriculum activity. (For more detailed descriptions of SAIL, see Bergman & Schuder, 1992; Pressley, El-Dinary, et al., 1992; Schuder, 1993.)

One hypothesis evaluated here was that participating in SAIL would enhance reading comprehension as measured by a standardized test. A second was that there would be clear indications after a year of SAIL instruction of students learning and using strategies. A third was that students would develop deeper, more personalized and interpretive understandings of text after a year of SAIL.

These hypotheses were evaluated with low-achieving second-grade students, a group targeted by SAIL: SAIL was designed originally for introduction to elementary students in either first or second grade who were at risk for reading failure. It is intended as a dramatically richer and more engaging form of instruction than the skill-and-drill approaches so often delivered to at-risk students (Allington, 1991). Thus, the evaluation reported here involved contrasting the achievement of low-achieving second-grade students who participated in SAIL with that of five matched

groups of second-grade students receiving high quality, but more conventional, reading instruction.

Method

Participants

Teachers. The five transactional strategies instruction teachers served in the school district that developed the SAIL program; the five teachers in comparison classrooms were from the same school district. This district had garnered numerous national awards for excellence in instruction. Eight of the teachers taught second-grade classes. One SAIL teacher had a first/second-grade combination; one comparison teacher had a second/third-grade combination. All teachers were female. The SAIL teachers had 10.4 years of experience teaching on average; the comparison teachers averaged 23.4 years on average.[1] The five SAIL teachers exhausted the pool of second-grade teachers in the district with extensive experience teaching in the SAIL program (i.e., 3 or more years; range = 3–6 years). The comparison teachers were recommended by principals and district reading specialists, with nominations of effective second-grade teachers made on the basis of criteria such as the following: (a) They gave students grade-level-appropriate tasks; (b) they provided motivating learning activities; (c) they used classroom management well to avoid discipline problems; (d) they fostered active student involvement in reading; (e) they monitored student understanding and performance; and (f) they fostered academic self-esteem in students.

The comparison teachers were eclectic in their instructional practices, blending the whole-language tradition favored in the school district with elements of skill and other traditional forms of conventional reading instruction. For example, a teacher who stressed skills instruction sometimes integrated literature-based activities such as having students write in a response journal or read from a trade book (rather than a basal reader). A teacher who emphasized elements associated with a literature-based approach also taught or reviewed phonics, word attack, and specific comprehension skills before or after reading. Some conventional instructors also taught a few strategies, like skipping unknown words, making predictions, or activating background knowledge. However, they did not teach a flexible repertoire of strategies using explicit, verbal explanations of thinking, elements characteristic of SAIL instruction. The comparison-group teachers had not participated in any SAIL professional-development activities.

All participating teachers were administered DeFord's (1985) Theoretical Orientation to Reading Profile (TORP), a 28-item instrument discriminating among teachers identifying with phonics, skills, and whole-language orientations. The scoring is such that those favoring phonics-based reading instruction score lower than those favoring skills instruction, who score lower than those identifying with whole language (scores range from 28 to 140). The SAIL teachers' mean score was 113 (SD = 9.7), and the comparison teachers averaged 73 (SD = 7.2), with the SAIL teachers differing significantly from the comparison teachers, dependent $t(4) = 6.24, p < .05$. (In the teacher comparisons, teachers were the units of analysis. Each teacher taught a reading group, with each group consisting of six students. SAIL and non-SAIL reading groups were matched on school demographic information. SAIL and non-SAIL students in the matched reading groups were paired on the basis of students' fall standardized test performances, described later in the Method section. As such, a correlated-samples analysis was conducted, because SAIL and comparison teachers were matched.) When the particular items of the TORP were examined, it was clear that the SAIL teachers had more of a

whole-language orientation than the comparison teachers, who endorsed phonics and skills to a greater degree, smallest $|t|(4) = 4.88, p < .05$ for any of the three subscales. This finding was as expected, because SAIL encourages meaning-making as the goal of reading and discourages teaching of skills in isolation, consistent with whole language. Informal observations of the comparison teachers over the year confirmed that they were more eclectic in their approach to reading instruction than the SAIL teachers, incorporating a balance of whole-language, phonics-based, and skills-based instruction. Thus, their more balanced appraisal of the TORP items was consistent with our observations of their teaching.

At the beginning of the study, the 10 participating teachers were also administered a 25-item researcher-constructed questionnaire tapping their beliefs about teaching (r = .94; Cronbach's alpha calculated on participating teachers' responses). The questions involved responding to Likert-type statements (i.e., on a *strongly agree* to *strongly disagree* scale). For example, teachers who endorse transactional strategies instruction were expected to respond affirmatively to "The most important message to convey to students is that reading and thinking are inseparably linked," and "During instruction, teachers should ask story-related questions that have no precisely 'right' or 'wrong' answer." It was expected that SAIL teachers would disagree with items such as "Worksheets that enable students to practice comprehension skills can be very useful for low-group students," and "During reading instruction, teachers need to guide students towards one best interpretation of a story." The responses were scored so that consistency with transactional strategies instruction would result in a low score (maximum score = 120; one item was discarded). The scores of the SAIL teachers ranged from 25 to 45 on this scale (M = 36.8, SD = 9.5), and comparison teachers' scores ranged from 62 to 76 (M = 70.8, SD = 5.3), a significant difference, dependent $t(4) = -8.84, p < .05$. In short, there were multiple indicators at the outset of the study that the SAIL teachers were committed to a different approach to teaching from the conventional teachers and that the SAIL teachers' beliefs about teaching were consistent with a transactional strategies instruction philosophy.

Students. Student participants were assigned to second grade but were reading below a second-grade level at the beginning of the year. They were identified as such through informal testing (teacher assessments involving reading of graded basal passages and word lists), results from assessments administered as part of the Chapter 1 program, and the previous year's grades and reports. Unfortunately, none of the assessments used by the school district to classify readers as weak at the beginning of the year were standardized measurements, although there was converging evidence from the informal measures that all participants experienced

---

[1] We recognize that some readers may be concerned about the mean difference in years of teaching between the SAIL and non-SAIL teachers. In this study, the SAIL and non-SAIL classes were matched as closely as possible. The primary criteria for matching classes were demographic in nature. To the extent that it was possible, we used student mobility patterns, Chapter 1 status, ethnic and minority composition, size and location of schools, and standardized test performances. At the time, years of teaching experience did not seem to be as critical as some of the other factors. Given our decision, there is no way to separate out the effect that years of experience may have had on the way teachers taught their students. However, readers should bear in mind that the comparison teachers were highly regarded for their teaching abilities by district personnel; therefore, if anything, their greater number of years of experience could be construed as an advantage.

at least some difficulty reading beginning-level, second-grade material. Therefore, student mobility patterns, Chapter 1 status, ethnic and minority composition, size and location of schools, and overall performances on standardized tests were used to pair SAIL and comparison classes in the study. Moreover, because we did not have information about students' performance in previous years in any subject area and no formal test data existed for all these students, we administered a standardized achievement test. To attain greater comparability, a standardized achievement test was used to match students in each class as participants. A comprehension subtest of the Stanford Achievement Test (Primary 1, Form J; Grade level 1.5–2.5) was administered in late November or early December (depending on the class) of the school year. Administration of this test occurred at that point because only then did the teachers feel that participating students could function somewhat independently at the 1.5 grade level and thus not perform on the very floor of this measure. Unfortunately, this necessitated that the test be administered after SAIL teachers had introduced SAIL component strategies, so that it is not perfectly accurate to consider this a pretest.

Six students in each of the paired classes (i.e., a pair consisted of one SAIL class and one comparison class) were matched on the basis of their reading comprehension scores. All of the children participating in the study spoke and comprehended English. In addition, the sample included no children experiencing severe attentional or behavioral problems. Only six students in one SAIL class met the eligibility requirements. Because students were matched on the basis of their standardized comprehension pretest scores, six matched pairs were selected for participation. Between the first and second semesters, 1 SAIL student and 2 comparison students in one pair of classrooms left their classrooms. Backup students were substituted, with no significant difference occurring between the newly constituted groups on the fall reading comprehension subtest.

There were five reading groups for the SAIL condition and five groups for the non-SAIL condition, each consisting of 6 students per group. Thus, in all comparisons between conditions, the reading group mean was the unit of analysis, with each unit consisting of the mean of the 6 designated students in each group.[2] With a maximum raw test score of 40 possible, the SAIL classes in the study averaged 22.20 on this measure ($SD = 6.85$) at the late fall testing, and the comparison classes averaged 22.67 ($SD = 5.89$), a nonsignificant difference (means per class analyzed), $t(4) = -0.59, p > .05$. Although not used for matching, the word skills subtest from the same standardized instrument was also administered (maximum score $= 36$ for the subtest), SAIL $M = 20.97$ ($SD = 2.76$), and comparison $M = 21.10$ ($SD = 3.40$), $t(4) = -0.10, p > .05$. The comparability of the paired groups is reflected well by considering their means and standard deviations on the fall Stanford Reading Comprehension subtest (see Table 1).

Although the 6 children from each classroom are referred to here as a reading group, their instruction varied through the year. First, reading was most often taught in homogeneous groups, although it also occurred during individualized and whole-class instruction. Second, participants often, but not always, remained members of the same homogeneous group over the course of the year (students who made great progress became members of another group). Because the SAIL program was offered to all children in the SAIL classrooms and the instruction in comparison classrooms did not resemble SAIL instruction, variable grouping did not pose a problem with respect to fidelity of treatment. The six participating children in each classroom did meet as a homogenous group for lessons that were formally analyzed, however. Even so, our use of the term reading group implies only that the 6 targeted children

received either SAIL or conventional instruction daily, always within their classrooms, and frequently in small groups of students.

Design

This was an academic-year-long quasi-experimental study, carried out in 1991–1992. The reading achievement of five reading groups of low-achieving second-grade students receiving SAIL instruction was compared to the reading achievement of five groups of low-achieving second-grade students receiving instruction typical of second-grades in the district. Each of the 10 reading groups was housed in a different classroom, with each SAIL group matched with a comparison reading group that was close in reading achievement level at the beginning of the study and from a school demographically similar to the school representing the SAIL group. That is, there were five matched pairs of reading groups (6 low-achieving students per reading group), with one SAIL and one comparison reading group per pair.

The present study incorporated a quasi-experimental design in that we did not randomly assign teachers to conditions. Preparing teachers to become competent transactional strategies instructors is a long-term process; therefore, we felt we could not randomly assign teachers, provide professional development, and wait for teachers to become experienced in teaching SAIL in a realistic time frame. Also, the sample incorporated the largest cohort of experienced SAIL teachers in the school system. Therefore, we decided not to take SAIL teachers and randomly assign one group to teach SAIL and one group to teach another method. Even if we had access to a larger pool of SAIL teachers, we would not have asked them to alter for an entire year practices they were committed to.

The fact that SAIL teachers were committed to strategies instruction was not a concern; we felt that effective comparison teachers would be committed to the teaching practices they espoused as well. Although we might have attempted to identify potential comparison teachers in the buildings where SAIL teachers taught and randomly assigned students to teachers, we opted not to do this in favor of seeking the most competent second-grade comparison teachers that we could in the district. Because the comparison second-grade teachers did not serve in the same buildings as the SAIL teachers, random assignment of children to teachers was impossible. We believe the option we selected of matching reading groups taught by SAIL teachers with groups taught by teachers believed by the district administrators and reading consultants to be excellent second-grade reading teachers was a fair test of SAIL relative to highly regarded, more conventional reading instruction. We recognize that the use of a quasi-experimental design invites alternative explanations for results. However, we designed a study that was as close to experimental as possible by instituting as many precautions as we could.

Dependent Measures

The dependent measures are described in the order in which they were administered in the academic year. A summary of the measures appears in Table 2.

---

[2] All class means were based on 6 students with the exception of the following, which reduced this number because of either data loss or absence: Strategies interviews: 1 student in two non-SAIL classes (pretest), 1 student in one SAIL class (posttest); retellings: 1 student in one SAIL class ("Mushroom" story), 1 student in one SAIL class ("Fox Trot" story); think-aloud task: 2 students in one SAIL class.

Table 1

Stanford Achievement Test Scores: Matched SAIL and Non-SAIL Class Means and Standard Deviations For Word Attack and Reading Comprehension

| | Pretest | | | | | Posttest | | | | |
| | SAIL | | Non-SAIL | | | SAIL | | Non-SAIL | | |
Matched classes	M	SD	M	SD	t	M	SD	M	SD	t
T1 and T6										
Word Study Skills	19.33	6.35	16.33	3.08		28.00	4.05	22.33	5.75	
Reading										
Comprehension	20.83	6.94	20.33	7.92		36.83	4.36	26.83	9.28	
T2 and T7										
Word Study Skills	20.83	5.27	19.67	5.65		27.83	4.26	24.67	5.24	
Reading										
Comprehension	19.67	6.47	20.00	8.07		33.00	6.26	26.67	7.81	
T3 and T8										
Word Study Skills	18.67	4.72	23.33	2.73		23.67	2.80	22.83	6.01	
Reading										
Comprehension	15.67	4.63	16.83	6.24		30.67	1.63	26.00	9.94	
T4 and T9										
Word Study Skills	20.33	7.42	21.00	4.47		26.50	2.43	24.00	5.73	
Reading										
Comprehension	21.00	6.23	24.00	6.60		33.67	3.72	29.00	8.12	
T5 and T10										
Word Study Skills	25.67	2.50	25.17	5.04		29.50	5.79	26.17	5.60	
Reading										
Comprehension	33.83	7.28	32.17	6.88		36.83	2.40	35.17	4.22	
Group totals										
Word Study Skills	20.97	2.76	21.10	3.40	-0.10	27.10	2.19	24.00	1.53	3.98*
Reading										
Comprehension	22.20	6.85	22.67	5.89	-0.59	34.20	2.65	28.73	3.77	4.02*

Note. SAIL = Students Achieving Independent Learning. Maximum possible score on Word Study Skills subtest was 36, and on Reading Comprehension subtest, 40. SAIL and non-SAIL differences on Word Study Skills and Reading Comprehension pretests tested at α = .05 (one-tailed).
* $p < .05$. one-tailed.

Strategies interview. In October and November (i.e., when SAIL components were being introduced to SAIL students) and in March and April, a strategies interview was administered to all students participating in the study. This interview tapped students' reported awareness of strategies, as measured by the number and types of strategies they claimed to use during reading. We also hoped to assess whether students were aware of where, when, and why to use strategies.

Five open-ended questions (adapted from ones used by Duffy et al., 1987, for their study of strategies instruction with third-grade readers) were administered orally and individually to each participating student:

1. What do good readers do? What makes someone a good reader?
2. What things do you *do* before you start to read a story?
3. What do you *think* about before you read a new story?
4. What do you do when you come to a word you do not know?
5. What do you do when you read something that does not make sense?

These questions were presented in a different order for each student. If initial student responses were unclear or terse, the researcher probed for clarifications and elaborations.

Story lessons and retelling questions. In March or April (depending on class schedule), two stories were read by all participating reading groups. The instruction and interactions that occurred during reading were recorded as these stories were read, and they were analyzed to document differences in instruction in the

SAIL and non-SAIL reading groups. (See the Appendix for a description of two SAIL and two non-SAIL lessons serving as a general comparison of SAIL and conventional group instruction.) A descriptive analysis of the lessons revealed that SAIL teachers more often gave explicit explanations, verbalized their thinking, and elaborated explicitly and responsively in reaction to students' comments and actions. Non-SAIL teachers more frequently than SAIL teachers provided information or instruction to students without stating a purpose, gave answers to students when they had difficulty reading or answering questions, drilled students on their learning, and praised and evaluated their performance. Both groups activated students' background knowledge, reviewed previously learned information, and guided students through their difficulties to about the same extent (Brown, 1995a, 1995b).

After the lesson was conducted, each student was asked to retell the story to a researcher, followed by a task requiring students to sequence pictures corresponding to events in the story. The primary purpose of this measure was to assess students' recall of text details, although we thought students might include interpretations in their retellings of story content as well.

All reading groups in the study read the same two illustrated stories. "Fox Trot" was a chapter in a popular children's trade book, *Fox in Love* (Marshall, 1982); "Mushroom in the Rain" (Ginsburg, 1991) was from the Heath Reading Series, Book Level 1. The readability for the 341-word "Fox Trot" was 2.4; the readability for the 512-word "Mushroom in the Rain" was 2.2 (Harris–Jacobson Wide Range Readability Formula; Harris & Sipay, 1985, pp. 656–673).

Table 2
Description of Data Sources for Students

Data source	When given	Why given	Description
Strategies interview	October–November March–April	To assess SAIL and non-SAIL classes' awareness of comprehension and problem-solving strategies	Semistructured interview consisting of five base questions that were followed up with nondirective prompts; the questions were administered orally and individually to students.
Retelling questions	March–April	To assess students' retelling and sequencing of two stories presented by each teacher	Students individually were asked cued and picture-cued retelling questions.
Think-aloud task	May–June	To compare SAIL and non-SAIL classes' independent use of strategies during story reading; to determine if students were more text- or reader-based in their responses to probes	Students were stopped and asked "What are you thinking?" and other nondirective follow-up probes at four fixed points during story reading; students were questioned individually.
Standardized subtests of reading comprehension and word skills	November–December	To form comparable SAIL and non-SAIL reading groups by matching students using Stanford Achievement Test Reading Comprehension scores (fall administration)	Stanford Achievement Test, Primary 1, Form J
	May–June	To compare SAIL and non-SAIL classes on traditional, standardized, and validated measures of reading (fall and spring administration)	Stanford Achievement Test, Primary 1, Form K

Note. SAIL = Students Achieving Independent Learning.

In "Fox Trot," the main character, Fox, decides to enter a dance contest. He asks two friends to be his dance partner, but they refuse. They suggest that Fox ask Raisin, but Fox is reluctant to do so because Raisin is angry with him. Nevertheless, he asks and she agrees. They practice hard and dance quite well together. On the day of the contest, Raisin gets the mumps. Fox returns home and despondently sits in front of a blank TV. Then he decides to teach his sister the dance steps. They rush to the contest and claim second prize.

In "Mushroom in the Rain," an ant seeks shelter from a storm. She squeezes herself into a small mushroom. A butterfly comes by and asks if he can escape the rain as well, with the ant allowing the butterfly to crowd in. Then comes a mouse and a bird, with the crowding in the mushroom increasing. A rabbit then arrives, who is being chased by a fox. The others hide the rabbit in the mushroom. Once the fox leaves and the rain stops, the ant asks the others how they managed to fit under the mushroom. A frog, sitting on top of the mushroom asks, "Don't you know what happens to a mushroom in the rain?" In the version of the story used in the study, the answer was not provided to the children but was left for them to infer.

These stories were selected because they provided ample opportunity for diverse interpretations and personal responses. They were on the school system's approved list and approved by the participating teachers as appropriate for a single lesson for weaker second-grade students in the spring.

All decisions about how to present the stories were made by the teachers. However, they were asked to present each of these stories in one morning lesson that was not to exceed 55 min in length. They were consistent with this request, with the mean SAIL lesson lasting 43.40 min (SD = 7.83) and the mean comparison-group lesson lasting 35.50 min (SD = 13.34). (Three of the matched pairs of reading groups read "Mushroom in the Rain" first; two pairs

read "Fox Trot" first.) Generally, SAIL lessons are lengthier because negotiating interpretations, explaining and modeling strategies, thinking aloud, and selecting and using "fix-up" strategies while reading are time-consuming activities, particularly when they are compared with some activities in conventional reading lessons (e.g., answering skill-and-drill and literal comprehension questions). The lessons were videotaped to allow a manipulation check to ascertain that teaching in the SAIL groups was different as expected from teaching in the comparison reading groups (described in the Results section).

Approximately 2 hr after each lesson was over, each of the 6 students in the reading group was interviewed individually. First, students were asked to retell the story:

Pretend that you are asked to tell the story to other kids in the class who have never heard the story before. What would you tell them happened in that story? Can you remember anything else? (Adapted from Golden, 1988)

This interview was followed by a cued, picture retelling task. Students were asked to sequence six scrambled pictures taken directly from the story. The students were then informed that sometimes pictures assist in aiding recall of stories, and they were asked to use the pictures to prompt recollection of story content.

Think-aloud measure. In May or June, students read a 129-word illustrated Aesop's fable, "The Dog and His Reflection," selected from a trade book (Miller, 1976). The readability for this story was 3.9 (Harris & Sipay, 1985), making it challenging for the students.

In the story, a dog steals a piece of meat from the dinner table. He runs into the woods and starts to cross a bridge. When he chances to look down, he sees his reflection in the water. Thinking his reflection is another dog with a larger cut of meat, he decides

to seize the dog's chop. When he opens his mouth, his own piece of meat plunges into the water. Consequently, the dog ends up with nothing at all.

The students met with the researcher individually for this task. Students were stopped at four points in the reading of the Aesop fable and asked to report their thinking. If a student had difficulty reading a segment, the first question posed was, "What do you think happened on this page?" Otherwise, the student was asked first, "What are you thinking?" Both questions primarily focused on content, with the "What are you thinking?" probe designed to be open-ended enough to elicit interpretive remarks and opinions about the fable, although we expected students to recount story details as well. Thus, the first purpose of the think-aloud task was to supplement the story-recall task. Unlike the recall questions that were designed primarily to assess memory for story details, the more open-ended, think-aloud prompt was used to examine students' understandings and interpretations of text.

The other purpose of this measure was to supplement the strategies interview. Although the strategies interview revealed whether students talked about strategies, it did not indicate whether students used them on their own when reading. One limitation of the strategy interview was that students might memorize information repeated by their teachers without being able to translate that knowledge into practice. Therefore, a task was designed to observe whether students actually used comprehension strategies when reading. Our intent was not to have these young students report directly on their thinking processes while reading. Instead, we observed whether students would use comprehension strategies when they were not cued to do so.

When students offered unelaborated responses to initial questions, open-ended follow-ups were asked (see Garner, 1988, p. 70), such as, "Can you tell me more?" or "Why do you say that?" Sometimes an unelaborated comment was echoed back to the student in the form of a question. Thus, after a student remarked that a dog stole a piece of meat from his master's table, the researcher asked, "What do you think about the fact that the dog stole a piece of meat from his master's table?" For every text segment, before the student moved on to reading the next segment, the researcher asked, "Is there anything you could say or do before reading on?"

Stanford Achievement Test subtests. In May or June, students took the Stanford Achievement Test (The Psychological Corporation, 1990), Form K, Reading Comprehension and Word Study Skills subtests. Standardized tests traditionally have been used as measures of reading performance in strategy experiments. Therefore, in addition to the other measures, students were compared on a conventional measure of reading achievement.

The Reading Comprehension subtest consists of two-sentence stories, comprehension questions on short passages, and sentence-completion items that form short stories. The Word Skills subtest includes questions pertaining to structural analysis (e.g., compound words, inflectional endings, contractions) and phonetic analysis (e.g., consonants and vowels). The comprehension test was administered first to all students, followed by the word skills test. The alternate-forms reliability for the full scale scores of Forms J (administered in the fall) and K was .89.

Results

Every hypothesis tested here was one-tailed, and each was an evaluation of whether SAIL instruction produced better performance than the comparison instruction. Most means appeared in only one hypothesis test, and hence, $\alpha <$

.05 was the Type 1 error probability selected for all hypotheses (Kirk, 1982, for this and all references to statistics). For the standardized test data and strategies interview data, the simple effect of condition within time of testing was evaluated in the fall, as it was in the spring. The Time of Testing × Condition interaction was also tested. The hypothesis-testing approach taken here was conservative, providing high power for detection of large effects only (Cohen, 1988). For each dependent variable, the same overall Type 1 error probability would have occurred if we had analyzed the data within a 2 × 2 analysis of variance structure.

All tests were based on the reading group mean as the unit of analysis (i.e., $n = 5$ groups for the SAIL condition, and $n = 5$ groups for the comparison condition, each consisting of 6 students per group), because individual scores within reading groups were not independent (see Footnote 2). Finally, all t tests were dependent t tests that were based on the 5 matched pairs, with one SAIL and one comparison group to a pair, with pairings determined by demographic information and by the reading groups' fall standardized comprehension performances, as described earlier.

For every dependent t test involving student posttest performance, an exact permutation test was also conducted. In all cases except three, performance in SAIL classes significantly exceeded performance in the comparison classrooms, $p = .03125$ (one-tailed) for the permutation test. In the two exceptions reported in the main text (i.e., the pretest-to-posttest gain on the standardized comprehension measure and the pretest-to-posttest gain on the strategies interview: word attack strategies), the gains for one of the 5 SAIL and non-SAIL pairs were identical. The SAIL classes exceeded the non-SAIL classes in the other pairs for both measures ($.03 < p < .07$, one-tailed). The third exception was in a supplementary analysis.[3]

In general, the results are reported in the order in which dependent measures were described in the Method section, which parallels the order of data collection in the study.

---

[3] One reviewer strongly felt that the skipping strategy was not as "good" or "useful" as some of the other strategies students reported using. Consequently, we are providing data so that readers can compare the two groups specifically on the skipping strategy. For the fall strategies interview, the SAIL sum of mean frequencies by group for skipping as a word attack strategy was 3.93 ($SD = 0.35$), and the non-SAIL summed mean was .94 ($SD = 0.70$), $t(4) = 8.20, p < .05$. For the spring strategies interview, the SAIL summed group mean for skipping as a word attack strategy was 4.33 ($SD = 1.10$), and the non-SAIL summed mean was 1.47 ($SD = 1.55$), $t(4) = 5.36, p < .05$. The interaction was $t(4) = -0.36, p > .05$. For the fall strategies interview, the SAIL sum of mean frequencies by group for skipping as a *comprehension strategy* (i.e., ignoring a larger segment of text and reading on) was .17 ($SD = 0.40$), and the non-SAIL summed mean was .57 ($SD = 0.50$), $t(4) = -1.10, p > .05$. For the spring strategies interview, the SAIL summed group mean frequency was 2.47 ($SD = 1.25$) and the non-SAIL mean frequency was .83 ($SD = 0.80$), $t(4) = 2.14, p < .05$. The interaction was also significant, $t(4) = 2.20$. (The permutation test for the interaction was not significant, however, because of one tie in the data.)

Fall–Spring Strategies Interview

The interviews were designed to determine whether SAIL and comparison students would differ in their awareness of strategies, operationalized as the number of strategies they claimed to use during reading. Two raters scored 20% of the interviews, with an overall 87% agreement for the strategies named by students. Only one of the two raters scored the remainder of the interviews.

A strategy was scored as mentioned if it was named in response to any of the interview questions. Any strategies mentioned by students were recorded, even if they were not strategies taught in the SAIL program. The comprehension strategies mentioned included the following:

Predicting: Guessing what will happen next

Verifying: Confirming that a prediction was accurate

Visualizing: Constructing a mental picture of the information contained in the text segment

Relating prior knowledge or personal experiences to text: Making an association between information in the text and information in the readers' head

Summarizing or retelling: Saying the most important information (summarizing) or restating in one's own words everything that occurred in the text segment just read

Thinking aloud: Verbalizing thoughts and feeling about text segments just read

Monitoring: Explicitly verbalizing when something just read does not make sense

Setting a goal: Deciding a purpose prior to reading, including decisions about both expository and narrative texts

Browsing or previewing: Flipping through the story, glancing at the pictures, or reading the back cover to get ideas about the story

Skipping: Ignoring a problematic part of text and reading on

Substituting or guessing: Replacing a difficult part of text with something else that appears to make sense and maintains the coherence of the text segment

Rereading: Returning to a problematic segment of text

Looking back: Looking back in the text for information that might help in understanding a difficult-to-understand part of text

Clarifying confusions: Asking a specific question to resolve a comprehension problem

Asking someone for help: Asking another student or the teacher for help with the confusing section of text

The following strategies for attacking unknown or difficult words were mentioned:

Skipping: Ignoring a problematic word and reading on

Substituting or guessing: Replacing an unknown word with another word that appears to make sense or that maintains the coherence of the text segment

Rereading: Returning to a problematic word

Looking back: Looking back in the text for information that might help in understanding a difficult-to-understand word

Using picture clues: Looking at pictures in the story to help determine the meaning of an unknown word or difficult-to-understand piece of text

Using word clues: Relying on the surrounding text to help decide the meaning of an unknown word or difficult-to-understand piece of text

Breaking a word into parts: Seeing if there are recognizable root words, prefixes, or suffixes contained within the larger word

Sounding out a word: Applying knowledge of phonics to the decoding of the word

Asking someone for help: Asking another student or the teacher for help with the confusing word

The comprehension and word-level strategies reports are summarized in Table 3. The means reported in the table are reading group means (i.e., a mean frequency of strategies reported for each reading group in the study was calculated on the basis of individual reading group members' reports, with each of the Table 3 means and standard deviations calculated on the basis of five reading group means). With respect to reports of comprehension strategies, there was no significant advantage for the SAIL students in the fall, shortly after the program had begun. By spring, however, as expected, the SAIL groups reported many more strategies than the comparison groups. In the spring, only SAIL students reported visualizing, looking back, verifying predictions, thinking aloud, summarizing, setting a goal, or browsing. Although during the spring interview, comparison-group students mentioned predicting, using text or picture clues to clarify confusions, making connections between text and their background knowledge and experiences, asking someone for help, skipping over confusing parts, and rereading, the mean frequency of such reports was always descriptively lower for them compared to the SAIL students. The SAIL and comparison groups mentioned monitoring and guessing approximately equally on the spring interview.

There were qualitative differences in students' responses to the strategy interview questions as well. When asked, "What do good readers do?" SAIL students responded more frequently than non-SAIL students that good readers use comprehension strategies, apply problem-solving strategies, and think. Both groups mentioned that good readers read abundantly, practice frequently, read well, and read for enjoyment. In response to questions about what students do or think before they read a story, students in both groups said they made predictions. However, SAIL students tended to predict what would happen in the story, whereas non-SAIL students predicted whether the story would be too difficult or whether they would like it. When asked, "What do you do when you read something that does not make sense?" students in both groups frequently mentioned they would skip or reread a confusing section; however, SAIL students cited these strategies more frequently.

With respect to word-level strategies, the SAIL students reported more strategies than the comparison-group participants, even during the fall interview (see Table 3). In the fall, SAIL students mentioned skipping words (see Footnote 3), substituting or guessing, using picture or word clues, rereading, and breaking words into parts descriptively more often than did comparison students. There was slightly more mention of sounding out of words in the comparison condition in the fall. The introduction to SAIL from the very start of school probably accounts for this fall difference in word-level strategies reports. By the spring, all of the word-

Table 3
Means and Standard Deviations for Number of Comprehension and Word-Level Strategies Mentioned in the Fall and Spring Strategies Interviews

| | Fall | | | | | Spring | | | | |
| | SAIL | | Comparison group | | | SAIL | | Comparison group | | |
Strategy	M	SD	M	SD	t(4)	M	SD	M	SD	t(4)
Comprehension	0.79	0.45	0.88	0.44	0.58	4.20	0.86	1.25	0.48	9.53
Word level	2.16	0.79	1.15	0.28	3.52	3.22	0.63	1.68	0.37	4.83

Note. SAIL = Students Achieving Independent Learning. With the exception of comprehension data in fall interviews, SAIL data were significantly greater than comparison data, $p < .05$, one-tailed.

level strategies were being mentioned by SAIL students. In contrast, the only word-level strategies mentioned consistently by more than 1 student per comparison reading group were skipping an unknown word, sounding a word out, rereading, and asking someone for help.

We also tested whether SAIL students made greater gains in self-reported awareness of strategies over the course of the year. The one-tailed interaction hypothesis test (e.g., fall-to-spring increase in students' strategies scores by condition) was significant, as expected, for both the comprehension strategies, $t(4) = 7.72$, and the word-level strategies, $t(4) = 2.64$.

In general, SAIL students provided more elaborate responses to postmeasure questions. For example, this rich spring interview was provided by a student in a SAIL class:

R: What do good readers do?
S: [Good readers have] lots of expression. They do think-alouds.
R: They do think-alouds. Okay. What do you mean by that?
S: Well, they tell people what they think is going on in their own words in the story.
R: Uh huh. What other things do good readers do?
S: Well, I'm an expert reader. And what I do is I skip. But, well, skipping isn't always great because sometimes you need to get the gist of the story. Cause if you always skip, you can't get the meaning of the story.
R: So you can't be skipping everything in the story . . . .
S: [I also do] substituting, and sounding things out is a very good strategy [sic] . . . and, um, looking back is a good strategy.
R: Looking back . . . . Why is looking back a good strategy?
S: Because like if I got stuck on a word, like, uh, it might be back on the story . . . . But sometimes it isn't.
R: Are there any other things good readers do? Are there any other strategies good readers use?
S: Guessing too. Picture clues are very good . . . . "The Cat and the Canary" has beautiful illustrations, and we think it should have a Caldecott medal 'cause the picture clues. I looked there and the word "suddenly" came up because the picture clues just looked like "suddenly."
R: What things do you *do* before you start to read a story?
S: I look at the title, and I look first at the pictures.
R: Why do you do that?
S: Because that can give me information about what the story is about. But when I make predictions it's not always right. We don't get upset because it's not right. We just know that it's not right and then something goes

off in our mind telling us that we should make another one.
R: What do you *think* about before you read a new story?
S: I think about whether it might be good or bad . . . .
R: How would you tell if it were good or bad?
S: If I were alone at home, I would look at the first pictures and start reading the first page and then I get ideas.
R: Okay, then, what might you do after you read the first page and get ideas?
S: I have a think-aloud in my mind that would tell me what the story might be about.
R: What do you do when you come to a word you do not know?
S: I use picture clues, I guess, look back, and sometimes I reread the sentence.
R: What do you do when you read something that does not make sense?
S: I read the sentence very slowly to see if I skipped a word.
R: Hmmm, what else do you do?
S: Sometimes I just skip it and go to the next line.

The following interview is representative of the type of responses given by non-SAIL students to the interview questions. Although some of the same components are apparent (particularly with word-level strategies), the student's responses are less elaborated:

R: What do good readers do?
S: They read a lot of books.
R: Anything else?
S: Nope.
R: What things do you *do* before you start to read a story?
S: Read the title.
R: Read the title. Why do you read the title?
S: Because, when, . . . if you don't read the title you won't know what it's about.
R: What do you *think* about before you read a story?
S: It might be tales.
R: It might be tales . . . what do you mean? Tell me a little more. . . .
S: Like, it might be funny.
R: Ah, so it might be funny . . . and how might you find that out? You haven't started reading it yet.
S: You might ask someone who read the story.
R: And what do you do when you come to a word you do not know?
S: You could ask your mother.
R: Is there anything else you could do?
S: I skip and then read the other words and then when you

have finished the sentence, you could go back to that letter and you can sound it what it is.

R: What do you do when you read something that does not make sense?

S: You might read the word that you don't know and you're not sure what it is.

R: Anything else?

S: No.

Although the SAIL students mentioned a descriptively greater number of comprehension and problem-solving strategies than non-SAIL students, their responses did not reflect a high degree of complex reasoning about why using strategies is so beneficial. Students exhibited some rudimentary knowledge of when to use strategies appropriately: they were able to respond to questions about what they did or thought before reading and when encountering problems. Also, students were starting to understand that strategies could be used flexibly, especially for problem solving. Mentioning several strategies may have suggested some prerequisite understanding about the adaptive use of strategies. However, students' responses typically did not indicate precise conditions under which certain strategies could be applied effectively.

In summary, by spring the SAIL students definitely reported more comprehension and word-level strategies during the interview than did comparison-group students. That SAIL students were already reporting more word-level strategies in the fall than comparison students probably reflected the effects of the first month or two of instruction in the program. By spring, every strategy except two was mentioned descriptively more often in the SAIL than in the comparison group. The exceptions were sounding it out (which was consistent with the teaching philosophy of the comparison teachers) and asking for help with a word (which is difficult to construe as a strategy associated with independence in reading). Most important, SAIL students learned more about comprehension and word-level strategies over the year than comparison students. However, in general, this information concentrated more on awareness and naming of strategies than on deep understanding of how strategic reasoning works. The results suggest that fully self-regulated thinking is the product of years of development. Perhaps, too, the questions were neither precise nor concrete enough to probe the understanding of young children in an in-depth manner. Furthermore, the students may not have been able to verbalize knowledge of their own strategic processing (Pressley & Afflerbach, 1995).

Spring Story Lessons

Teaching of the lessons. The March–April lessons were transcribed from the videotape records, with the transcriptions read by four raters who were "blind" to condition.[4] One rater was a SAIL program developer, and the other three were graduate students familiar with transactional strategies instruction and the SAIL program in particular. The program developer correctly classified 9 of the 10 SAIL lessons as consistent with the intent and original vision of

the SAIL program; this rater definitely was sensitive to whether teachers explained and modeled strategic processes and encouraged interpretive construction of text meaning by students through use of comprehension strategies. The curriculum developer looked for evidence that the teachers thought aloud in their lessons and coached students to engage text actively (i.e., to relate text content to prior knowledge as well as to apply other strategies as appropriate). He classified all of the comparison lessons as not consistent with the SAIL approach and, in fact, not even close to being consistent with SAIL. The three graduate students correctly classified lessons as SAIL or non-SAIL for 59 of the 60 ratings made. Thus, there were clear instructional differences between the SAIL and non-SAIL classrooms during the March–April lessons.

Two raters reviewed the lessons (one rater was "blind" to condition) for evidence of strategies teaching, with interrater agreement of 85% and disagreements resolved by discussion (see Footnote 4). Collapsing across the two lessons observed for each teacher, a mean of 9.20 ($SD = 1.92$) different comprehension strategies were observed in the SAIL lessons compared to a mean of 2.00 ($SD = 0.71$) in the comparison lessons, $t(4) = 7.43$. Predicting, relating text to background knowledge, summarizing, and thinking aloud were observed in all SAIL groups. Only relating to background knowledge was observed in all comparison groups. In no SAIL group were fewer than seven of the comprehension strategies taught; in no comparison group were more than three observed.

On average, again collapsing across each participating reading groups' two lessons, 4.80 ($SD = 0.45$) word-level strategies were observed in the SAIL groups, and 4.00 ($SD = 0.71$) were documented in the comparison reading groups, $t(4) = 4.00$. Using semantic context clues and using picture clues were observed in all SAIL groups; using picture clues and sounding words out were observed in all comparison classrooms. The range of word-level strategies was between 4 and 5 in the SAIL groups and between 3 and 5 in the comparison groups. Thus, one important indicator that the instruction in the SAIL groups differed from comparison instruction was that there was more strategies in-

---

[4] We recognize that to rule out possible alternative explanations of the results, the two raters conducting interrater agreement should be "blind." However, there is a perspective held by some qualitative researchers that the use of blind raters does not do justice to the analysis of data because the blind rater has spent so little time immersed in the experiences that have led to the primary researcher's breadth of understanding. Thus, "expecting another investigator to have the same insight from a limited data base is unrealistic" (Morse, 1994, p. 231). We concurred to some extent with this argument; however, in attempting to reconcile positions, we opted for only one rater to be blind. In that way, the blind rater could lend credibility to the nonblind researcher's interpretations. In attempting to strike a balance, the nonblind researcher often deferred to the blind rater's opinion when a stalemate was reached. Also, when the primary researcher was unsure how to interpret the data in the transcripts and protocols that were not subjected to interrater agreement, the "blind" rater frequently assisted in the coding of the questionable segment or unit.

struction in the SAIL groups. The difference was much more striking with respect to comprehension strategies, however.

Student recall of stories covered in lessons. The recall protocols were analyzed using a modified analytic induction approach (Goetz & LeCompte, 1984); that is, coding categories emerged from analysis of the data. However, identification of categories also was highly informed by the work of O'Flahavan (1989) and Eeds and Wells (1989). In this study, only the results of the literal and interpretive analyses are presented, because only they relate directly to the stated hypotheses.

The full categorization scheme and analysis can be found in the work of Brown (1995a, 1995b). Both "Fox Trot" and "Mushroom in the Rain" were parsed into idea units, a variant of the T unit (Hunt, 1965). Loosely defined, an idea unit is a segment of written or oral discourse that conveys meaning, consisting of a verb form with any associated subject, object, and modifiers. Length or grammatical structure does not determine whether a segment is coded as an idea unit; what counts is whether the unit is meaningful. Interrater agreement was calculated for 20% of the recalled stories (see Footnote 4). It was 89% for classification of the protocols into idea units of various types (e.g., literal, interpretive).

A first issue addressed was whether SAIL students recalled more interpretive idea units than comparison students. These remarks reflected students' relating of background knowledge to text. Interpretive ideas were not explicitly stated in the text or in the pictures but did not contradict information in the text or pictures. For instance, for the Mushroom story, "He wanted to be dry" was scored as an interpretive remark. (The text had said, "One day an ant was caught in the rain. 'Where can I hide?' he wondered. He saw a little mushroom peeking out of the ground in a clearing and he hid under it.") Also, the comment, "But they tricked him," was scored as an interpretive unit for the Mushroom story. (The corresponding text was, "How could a rabbit get in here? Don't you see there isn't any room," said the ant. The fox turned up his nose. He flicked his tail and ran off.") As a third example, one not corresponding to any specific part of the Mushroom story, the remark, "And it was the only place to keep him dry," was coded as an interpretive remark because it was a conclusion that did not contradict anything in the text.

For the Mushroom story, SAIL groups averaged 6.12 interpretive units per student (SD = 1.54), which exceeded the corresponding figure of 4.48 in the comparison groups (SD = 1.70), $t(4)$ = 2.99. For "Fox Trot," SAIL groups averaged 5.58 interpretive units per student (SD = 1.63), which exceeded the corresponding figure of 3.84 in the comparison groups (SD = 1.63), $t(4)$ = 2.97.

In the example below, a SAIL student interjected a personalized interpretation into his retelling of story events. Interpreting occurred even though the task was not designed to elicit such information. The text stated that the frog asked the other animals if they knew what happened to a mushroom when it rained. He then hopped away, laughing. The student recall included the following response:

S: In the story, um um, the frog was just laughing because it was a miracle that came true. And the frog was laughing, the frog was laughing at them. And then really really when he was talking he said, "Don't you know what happens when it rains over a mushroom? And they they didn't know. They thought it was just a miracle, and when it was getting bigger it looked like a sleeping cap. So I think it was going wider and wider, and afterward when the sun came out and the fox was like an evil spirit, it went away. Um, they came, they came right out, and the mushroom was so big they didn't know what happened.

After the retelling was over, the researcher, curious about the origins of the student's interpretation, asked why he thought the fox was an evil spirit. The student replied, "Because it's like you know, the movies. And once there's this evil spirit and it's dark and nothing happens right. And once you kill it, the evil spirit, or if it goes away, and then it turns back into a good life." Thus, the student used his personal knowledge accrued from viewing movies to generate a unique interpretation that entered into his retelling.

In addition to scoring interpretive recall, we evaluated literal recall of ideas represented either in the stories or in the accompanying pictures. For example, one idea unit represented explicitly in the Mushroom story was, "He hid under it." If the student recalled this idea unit or a paraphrase of it, the student was scored as having recalled the unit. In "Fox Trot," there was a picture of Carmen and Dexter looking through a window, watching Fox dance. One idea unit was scored as recalled if the student reported something like, "His friends were looking at him dance from the window."

For the Mushroom story, SAIL reading groups recalled an average of 17.64 (out of a maximum of 79) literal idea units per student (SD = 3.95), which did not exceed literal recall in the comparison groups, who averaged 15.82 units (SD = 1.31), $t(4)$ = 1.10. For "Fox Trot," however, SAIL recall (M = 12.26 out of a maximum of 59 units; SD = 2.72) exceeded comparison-group recall (M = 8.38, SD = 2.94), $t(4)$ = 2.60.

In summary, SAIL students were significantly more interpretive in their recalls than comparison students, consistent with our expectations. Even though the questions called for literal recall of story content, SAIL students were more interpretive. This result is consistent with the conclusion that an interpretive propensity is internalized by TSI students. There were not strong expectations about the literal recall of the stories on the basis of condition, for we recognized that the comparison teachers covered the literal content of stories very well in their lessons. Even so, the students in the SAIL groups recalled more literal information than students in the comparison groups, although the difference favoring the SAIL students was significant for only one story.

One explanation of the story-recall results is that the SAIL story lessons were longer on average than the comparison-group lessons. Our impression throughout the conduct of this study was that SAIL students take more time when reading orally, with teachers frequently interjecting explicit explanations, requesting think-alouds, and elaborat-

ing responsively. Thus, we believe that at least the increased interpretations in the SAIL condition were due more to how time was spent in the SAIL lesson than to amount of time per se, although the design of this study does not permit a definitive conclusion on this point.

Spring think-aloud analysis. The think-aloud protocols generated by each student in reaction to the Aesop's fable about the dog and his reflection were transcribed and analyzed using an analytic induction approach (Goetz & LeCompte, 1984). Two raters (one rater was "blind" to condition) read through all of the protocols, independently taking notes and identifying potential categories of reported reading processes (see footnote 4). Through negotiation, a tentative set of process categories were identified, and these were applied by both raters independently to two protocols, one from a SAIL student and one from a comparison-group student. The two raters then met and refined the categories in light of the difficulties experienced scoring these two protocols. The refined categorization was applied to another pair of protocols, again independently by both raters. The refined categorizations captured all of the processes represented in these protocols, and thus, this set of processes was used to code all of the think-aloud protocols.

A response with any indication of comprehension strategy use was coded as "strategy-based." For example, the following excerpt was coded as a strategies-based response:

(The student read the page about the dog rushing out of the house with the piece of meat. The student then started to talk before the researcher asked an initial probe.)

S: I think my prediction is coming out right. (verifying)
R: Why do you say that?
S: Cuz, cuz I see a bridge over there and water. (using picture clues)
R: Uh huh . . . .
S: And he ran out of the house without anybody seeing him. Like I said before . . . .
R: Okay, so you think your prediction is right and you're using, you were pointing to the pictures.
S: Yep.

The specific strategies used were also coded using the comprehension strategy definitions from the strategies interview, with 89% agreement between two raters on 20% of the protocols on these codings of specific strategies. The mean number of strategies evidenced by SAIL reading group members (averaging across all groups) was 6.93 (SD = 1.46). The corresponding comparison-group mean was 3.18 (SD = 1.06). The SAIL readers applied significantly more strategies during the think-aloud task than did the comparison-group students, $t(4) = 9.59, p < .05$. In fact, there was no overlap in the group means, with SAIL group means ranging from 5.00 to 8.67 strategies used per student, on average, and corresponding comparison-group means ranging from 2.00 to 4.83. All strategies that were scored, except for one (monitoring), were observed descriptively more frequently in the SAIL than in the comparison protocols. The strategies that occurred in the SAIL condition, from most to least frequent, were as follows: predicting, relating text to prior knowledge, thinking aloud, summariz-

ing, using picture clues, verifying, seeking clarification, monitoring, looking back, visualizing, and setting a goal. The corresponding order for the comparison condition was predicting, using picture clues, verifying, relating text to prior knowledge, monitoring, seeking clarification, thinking aloud, and looking back. No apparent visualizing, summarizing, and setting a goal were observed in the comparison-group think-alouds.

We also examined whether SAIL or comparison groups focused more on text- or reader-based information when they did not respond strategically. Responses not classified as strategies-based were coded as either "text-based" or "reader-based" (interrater agreement on 20% of the protocols for classifying text- or reader-based responses was 94%).

Text-based responses contained information explicitly stated or pictured in the story. For example, after reading the first text segment, a student responded to the initial probe:

R: Okay, what are you thinking?
S: The dog stole something.
R: Uh huh . . . tell me more.
S: He knocked over the table.
R: He knocked over, talk nice and loud . . . he knocked things off the table . . . okay.
S: Yeah, and nothing really else.
R: Okay. And what do you think about what the dog did?
S: What do you mean?
R: What do you think about what the dog did?
S: He stole something.

Reader-based responses reflected a connection between the story and a student's prior knowledge, experiences, beliefs, or feelings. In the following example, a student read the segment about the dog stealing a piece of meat from the master's dinner table:

R: What are you thinking about what's happening on this page?
S: Sort of bad because I see that was part of their dinner, but they would not have all the uhm, protein.
R: Okay . . . .
S: The dog ate all that . . . .

Proportions were calculated for each class, indicating the relationship of text- and reader-based responses to the total number of responses that were not coded as strategies-based. From these class proportions, SAIL and comparison group means were computed. The mean for reader-based responding for the SAIL group was .74 (SD = .10). The mean proportion of reader-based responding for the comparison group was .45 (SD = 0.17). Thus, the SAIL group produced more reader-based responses than the comparison group, $t(4) = 3.61, p < .05$. Without exception, all SAIL classes were proportionally more interpretive than literal in their nonstrategies-based responses. In contrast, only 2 of 5 comparison classes were proportionally more interpretive in their responses.

In summary, the SAIL students used strategies on their own more than the comparison students. Although strategy use by itself does not constitute self-regulation, it does suggest that students had begun to apply strategies indepen-

dently, one aspect of self-regulated reading. Self-regulated readers are not only strategic; they also are goal-oriented, planful, and good comprehension monitors. Because we did not ask students to report directly on their strategic processing while reading, however, we cannot address those aspects.

In addition, the results of the think-aloud analysis supported the results of the recall analyses. For a story in which variable instructional time was not a factor, SAIL students made significantly more reader-based remarks than comparison students. The SAIL students responded more interpretively as well as personally.

Spring Standardized Test Performance

In May–June, the SAIL students outperformed the comparison students on the 40-item comprehension subtest. The reading group raw score mean in the SAIL condition was 34.20 (SD = 2.65); the corresponding comparison-group mean was 28.73 (SD = 3.77), $t(4)$ = 4.02 (see Table 1). The SAIL students also outperformed the comparison students on the 36-item word skills subtest, $t(4)$ = 3.98. The reading group word skills raw score mean in the SAIL condition in the spring was 27.10 (SD = 2.19); the corresponding comparison-group mean was 24.00 (SD = 1.53).

One of the most striking aspects of the spring comprehension standardized test data was the much lower variability among individual students within SAIL reading groups compared to comparison reading groups. (The careful matching of the reading groups in the fall was with respect to both mean performance *and* variability on standardized reading comprehension; thus, there was little difference in SAIL and comparison reading group variabilities in the fall, as reported in the Method section.) Also, with the exception of one pair of classes (T5 and T10), this lower variability among students in SAIL reading groups was evident in the spring word study skills data. This finding is obvious from examination of the standard deviations for each matched pair of reading groups on the standardized subtests (see Table 1).

We believed that an especially strong demonstration of the efficacy of the SAIL program would be greater gains on standardized measures over the course of the academic year in SAIL versus the comparison condition. Thus, we tested the size of the fall-to-spring increase in raw scores in the SAIL versus the comparison groups. The SAIL group averaged 22.20 on an alternate form of the comprehension subtest (SD = 6.85) at the late fall testing, indicating a fall-to-spring gain of 12.00 (SD = 5.20) on average, and the comparison classes averaged 22.67 (SD = 5.89) in the fall, yielding a fall-to-spring change of 6.07 (SD = 2.28) on average. For the word skills subtest, the fall SAIL mean was 20.97 (SD = 2.76), and the mean fall-to-spring increase was 6.13 (SD = 1.86). In the fall, the comparison mean was 21.10 (SD = 3.40), and the fall-to-spring mean difference was 2.90, (SD = 2.70). The one-tailed interaction hypothesis test was significant, as anticipated for the comprehension subtest, $t(4)$ = 3.70. The word skills subtest proved significant as well, $t(4)$ = 5.41.

In one of the matched pairs, there were some perfect scores on the comprehension posttest: The SAIL class mean was 36.83 (SD = 2.40); the non-SAIL class mean was 35.17, SD = 4.22). For this pair of reading groups, a version of the next level of the Stanford Comprehension subtest (Primary 2, Form J) was then administered. Consistent with the analyses reported in the last two paragraphs, the spring SAIL group mean was greater than the matched comparison-group mean, and the SAIL group standard deviation was lower than the comparison-group standard deviation: SAIL M = 29.8, SD = 5.42; comparison-group M = 21.8, SD = 10.17. (The pretest Reading Comprehension subtest mean for the SAIL class was 33.83 [SD = 7.28]; the mean for the non-SAIL class was 32.17 [SD = 6.88]).

In summary, by academic year's end, the SAIL second-grade students clearly outperformed the comparison-group students on standardized tests, with greater improvement on the standardized measures over the course of the academic year in the SAIL condition. Unfortunately, no additional end-of-year achievement data existed for the students for comparison purposes either in reading or in any other subject area.

On the standardized tests, gains in comprehension were expected because, more than anything else, SAIL is intended to increase students' understanding of text. The effects on students' word skills performance were more of a surprise, albeit a pleasant one, supportive of the SAIL intervention; we knew that all teachers, regardless of condition, taught phonics and word attack skills, although at different times of day (e.g., integrated into various content areas) and in different ways (e.g., covered in the form of worksheets or mini-lessons).

Discussion

We made many informal and formal observations throughout the 1991–1992 school year indicating that instruction in the SAIL and comparison classes was very different. The differences were apparent in the two lessons that were analyzed in the spring: A SAIL curriculum developer and several graduate students who were familiar with transactional strategies instruction had no difficulty discriminating between transcripts of SAIL and non-SAIL lessons. One important difference highlighted in the analysis of the spring lessons was that discussion of strategies was much more prominent in the SAIL than in the comparison reading groups. That the differences in teaching were so clear bolsters our confidence in this study as a valid assessment of the efficacy of SAIL with at-risk second-grade children.

SAIL had positive short-term and long-term impacts. In the short term, students acquired more information from stories read in reading group and developed a richer, more personalized understanding of the stories. Whether the focus is on the amount of literal information recalled from stories covered in reading group or student interpretations of the texts read, there were indications in these data of superior performance by SAIL students relative to the comparison students. We infer that SAIL students learn more daily

from their reading group lessons than do students receiving more conventional, second-grade reading instruction.

SAIL had long-term impacts as well. Consistent with our expectations, the SAIL students exhibited greater awareness of strategies by the end of the year than the comparison students. SAIL students also reported use of, or were inferred to use, strategies more than the comparison students: They thought aloud while reading the Aesop's fable at the end of the year. The standardized test performances of the SAIL students also were superior to those of the comparison students at the end of the year. Most critically, there was significantly greater improvement on standardized measures of reading comprehension from fall to spring in the SAIL versus the comparison classrooms. In short, all measurements of student reading achievement reported here converged on the conclusion that a year of SAIL instruction improves the reading of at-risk second-grade students more than does alternative high quality reading instruction.

This study is the strongest formal evidence to date that transactional strategies instruction improves the reading of elementary-level students. There were many elements taken into consideration in this study that varied freely in more informal comparisons of SAIL and alternative instruction, such as ones generated by the school district that developed the intervention: (a) The student participants were carefully matched in this investigation so that there was no striking difference in their standardized reading achievement at the outset of the study. (b) The teachers were carefully selected. From years of observing and interviewing committed SAIL teachers, we knew that they are excellent teachers in general, who offer rich language arts experiences for their students. Thus, it was imperative that a compelling evaluation of SAIL be in comparison to excellent second-grade instruction. Accordingly, we sought highly regarded comparison teachers. (c) The lessons analyzed in the transactional strategies instruction and comparison groups involved the groups' processing the same stories. (d) The same dependent measures were administered by the same tester so that measurement experiences were equivalent for participants.

Another strength of this evaluation was that it did not rely only on standardized assessments but included also assessments of students' reading that were grounded in their typical classroom experiences. The assessments of children's memories for and interpretations of stories read in class reflect better the day-to-day comprehension demands on students than do standardized measures. Although thinking-aloud measures are far from perfect indicators of thinking (Ericsson & Simon, 1980), the assessments of children's thinking as they read the Aesop's fable arguably tapped more directly the thinking processes of the children that SAIL was intended to change than did the standardized assessments.

Are the outcomes reported here generally significant beyond the specifics of the SAIL program? SAIL is a specific instantiation of reading comprehension strategies instruction as adapted by educators. Such instruction may serve as a model for other educators. SAIL provides teachers with a way to blend critical elements of direct teaching and holistic

principles of instruction, aspects of instruction that may already exist in conventional reading classrooms. Because many conventional programs already share features with SAIL (e.g., literature-based instruction, teaching of predicting and problem-solving strategies), these programs might be modified to include SAIL components.

As we argued at the beginning of this article, long-term, direct explanation of thinking processes and scaffolded practice of a manageable repertoire of powerful comprehension strategies constitute an approach replicated in a number of settings (see also Pressley, El-Dinary, et al., 1992, and Pressley & El-Dinary, 1993, for a number of examples). The practice has raced ahead of the science, however, with the educator-developed adaptations more ambitious in scope, more complex, and ultimately very different from the researcher-validated interventions (e.g., reciprocal teaching) that inspired the educator efforts. There is a real need to evaluate such adaptations, for there is no guarantee that the strategies instruction validated in basic research studies is effective once it is translated and transformed dramatically by educators.

The research reported here contrasts with basic research on strategies instruction in a number of ways. First, the intervention studied here was multicomponential and this study was not analytical at all with respect to components of the intervention. Typically, basic strategies instruction research has been much more analytical. We can defend this evaluation of an entire transactional strategies instruction package because the whole program is the unit of instruction in the schools we have been studying: When the interest is in whether an instructional package as a whole works, a study evaluating that whole relative to other instruction is definitely defensible, particularly if time spent in direct instructional activities is controlled carefully (e.g., in this study, both groups of students received a year of reading instruction in the context of a full year in the second grade). Moreover, it was not our intent to tease out which aspects of the program were most effective nor to determine which components in combination accounted for student gains, especially because we believe that the complex instruction exemplified by SAIL may be more than the sum of its component parts (Pressley, El-Dinary, et al., 1992).

Second, the program of research that includes this study is a mix of qualitative and quantitative research. In contrast, most basic studies of strategies have been quantitative only. We are certain that the quantitative study reported here would have been impossible without the 3 years of qualitative research leading up to it. At a minimum, that qualitative research affected the selection of dependent measures and the decision to study only accomplished SAIL teachers (see Pressley, Schuder, et al., 1992). More generally, it made obvious to us the scope of an investigation necessary to evaluate transactional strategies instruction so that the treatment would not be compromised by the evaluation.

Third, most basic strategies research is designed and conducted by researchers. When educators have participated in basic studies, it has been as delivery agents only. In the program of transactional strategies instruction research, researchers, program developers, and educators have com-

bined their talents to produce a body of research that realistically depicts transactional strategies instruction and evaluates it fairly. As the study was designed and as it unfolded, school-based educators were consulted frequently about the appropriateness of potential dependent measures and operations of the study. The result has been a much more complete and compelling set of descriptions of transactional strategies instruction and, now, a thorough appraisal of the impact of one transactional strategies instruction program on second-grade, weaker readers.

We do not claim that after 1 year of transactional strategies instruction these students have become self-regulated readers. Pressley and Afflerbach (1995) made the point that truly self-regulated reading is observed only in very mature readers. It has always been suggested that TSI needs to occur over the long term to be effective (Pressley, El-Dinary, et al., 1992). Our hypothesis is that true self-regulation is the product of years of literacy experiences, with TSI intended to get the process off to a good start. One year of such instruction at least gets second-grade readers who are experiencing difficulties in learning to read to improve their reading relative to a year for comparable students in very good conventional classrooms.

References

Allington, R. L. (1991). The legacy of "Slow it down and make it concrete." In J. Zutell & S. McCormick (Eds.), *Learner factors/teacher factors: Issues in literacy research and instruction: Fortieth yearbook of the National Reading Conference* (pp. 19–29). Chicago: National Reading Conference.

Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 353–394). New York: Longman.

Bell, R. Q. (1968). A reinterpretation of the direction of effects in studies of socialization. *Psychological Review, 75*, 81–95.

Bereiter, C., & Bird, M. (1985). Use of thinking aloud in identification and teaching of reading comprehension strategies. *Cognition and Instruction, 2*, 131–156.

Bergman, J., & Schuder, R. T. (1992). Teaching at-risk elementary school students to read strategically. *Educational Leadership, 50*(4), 19–23.

Brown, R. (1995a). *A quasi-experimental validation study of strategies-based instruction for low-achieving, primary-level readers.* Unpublished doctoral dissertation, University of Maryland at College Park.

Brown, R. (1995b, April). *The teaching practices of strategies-based and non-strategies-based teachers of reading.* Paper presented at the meeting of the American Education Research Association, San Francisco, CA.

Brown, R., & Coy-Ogan, L. (1993). The evolution of transactional strategies instruction in one teacher's classroom. *Elementary School Journal, 94*, 221–233.

Brown, R., & Pressley, M. (1994). Self-regulated reading and getting meaning from text: The transactional strategies instruction model and its ongoing evaluation. In D. Schunk & B. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and educational applications* (pp. 155–179). Hillsdale, NJ: Erlbaum.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Collins, C. (1991). Reading instruction that increases thinking abilities. *Journal of Reading, 34*, 510–516.

DeFord, D. (1985). Theoretical orientation to reading instruction. *Reading Research Quarterly, 20*, 351–367.

Dole, J. A., Duffy, G. G., Roehler, L. R., & Pearson, P. D. (1991). Moving from the old to the new: Research on reading comprehension instruction. *Review of Educational Research, 61*, 239–264.

Duffy, G. G., Roehler, L. R., Sivan, E., Rackliffe, G., Book, C., Meloth, M. S., Vavrus, L. G., Wesselman, R. Putnam, J., & Bassiri, D. (1987). Effects of explaining the reasoning associated with using reading strategies. *Reading Research Quarterly, 22*, 347–368.

Durkin, D. (1978–1979). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly, 12*, 481–538.

Eeds, M., & Wells, D. (1989). Grand conversations: An exploration of meaning construction in literature study groups. *Research in the Teaching of English, 23*, 4–29.

El-Dinary, P. B., & Schuder, R. T. (1993). Seven teachers' acceptance of transactional strategies instruction during their first year using it. *Elementary School Journal, 94*, 207–219.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*, 215–251.

Garner, R. (1988). Verbal report data on cognitive and metacognitive strategies. In C. E. Weinstein, E. T. Goetz, & P. A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 63–76). San Diego, CA: Academic Press.

Gaskins, I. W., Anderson, R. C., Pressley, M., Cunicelli, E. A., & Satlow, E. (1993). Six teachers' dialogue during cognitive process instruction. *Elementary School Journal, 93*, 277–304.

Ginsburg, M. (1991). Mushroom in the rain. In D. Alvermann, C. A. Bridge, B. A. Schmidt, L. W. Searfoss, P. Winograd, & S. G. Paris (Eds.), *My best bear hug* (pp. 144–154). Lexington, MA: Heath.

Goetz, J. P., & LeCompte, M. D. (1984). *Ethnography and qualitative design in educational research.* San Diego, CA: Academic Press.

Golden, J. M. (1988). The construction of a literary text in a story-reading lesson. In J. L. Green & J. O. Harker (Eds.), *Multiple perspective analyses of classroom discourse* (pp. 71–106). Norwood, NJ: Ablex.

Harris, A. J., & Sipay, E. R. (1985). *How to increase reading ability: A guide to developmental and remedial methods.* New York: Longman.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels* (NCTE Research Report No. 3). Champaign, IL: National Council of the Teachers of English.

Hutchins, E. (1991). The social organization of distributed cognition. In L. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 283–307). Washington, DC: American Psychological Association.

Johnston, P., & Afflerbach, P. (1985). The process of constructing main ideas from text. *Cognition and Instruction, 2*, 207–232.

Kirk, R. E. (1982). *Experimental design* (2nd ed.). Monterey, CA: Brooks/Cole.

Lytle, S. L. (1982). *Exploring comprehension style: A study of twelfth-grade readers' transactions with text.* Unpublished doctoral dissertation, University of Pennsylvania. (University Microfilms No. 82-27292).

Marks, M. B., Pressley, M., Coley, J. D., Craig, S., Gardner, R., DePinto, T., & Rose, W. (1993). Three teachers' adaptations of

reciprocal teaching in comparison to traditional reciprocal teaching. *Elementary School Journal, 94,* 267–283.

Marshall, E. (1982). *Fox in love.* New York: Penguin Books.

Mehan, H. (1979). *Learning lessons: Social organization in the classroom.* Cambridge, MA: Harvard University Press.

Meichenbaum, D. (1977). *Cognitive-behavior modification: An integrative approach.* New York: Plenum Press.

Miller, J. P. (1976). *Tales from Aesop.* New York: Random House.

Morse, J. M. (1994). Designing funded qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 220–235). Thousand Oaks, CA: Sage Publications.

O'Flahavan, J. F. (1989). *Second graders' social, intellectual, and affective development in varied group discussions about literature: An exploration of participation structure.* Unpublished doctoral dissertation, University of Illinois at Champaign-Urbana.

Olshavsky, J. E. (1976–1977). Reading as problem-solving: An investigation of strategies. *Reading Research Quarterly, 12,* 654–674.

Olson, G. M., Mack, R. L., & Duffy, S. A. (1981). Cognitive aspects of genre. *Poetics, 10,* 283–315.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1,* 117–175.

Paris, S., & Oka, E. R. (1986). Children's reading strategies, metacognition, and motivation. *Developmental Review, 6,* 25–56.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Hillsdale, NJ: Erlbaum.

Pressley, M., & El-Dinary, P. B. (Guest Eds.). (1993). Special issue on strategy instruction. *Elementary School Journal, 94*(2).

Pressley, M., El-Dinary, P. B., Gaskins, I., Schuder, T., Bergman, J. L., Almasi, J., & Brown, R. (1992). Beyond direct explanation: Transactional instruction of reading comprehension strategies. *Elementary School Journal, 92,* 513–555.

Pressley, M., Gaskins, I. W., Cunicelli, E. A., Burdick, N. J., Schaub-Matt, M., Lee, D. S., & Powell, N. (1991). Strategy instruction at Benchmark School: A faculty interview study. *Learning Disability Quarterly, 14,* 19–48.

Pressley, M., Goodchild, R., Fleet, J., Zajchowski, R., & Evans, E. D. (1989). The challenges of classroom strategy instruction. *Elementary School Journal, 89,* 301–342.

Pressley, M., Johnson, C. J., Symons, S., McGoldrick, J. A., & Kurita, J. A. (1989). Strategies that improve children's memory and comprehension of text. *Elementary School Journal, 90,* 3–32.

Pressley, M., Schuder, T., Teachers in the Students Achieving Independent Learning Program, Bergman, J. L., & El-Dinary, P. B. (1992). A researcher–educator collaborative interview study of transactional comprehension strategies instruction. *Journal of Educational Psychology, 84,* 231–246.

The Psychological Corporation (1990). *Stanford Achievement Test Series: Technical data report* (8th ed.). San Diego, CA: Harcourt Brace Jovanovich.

Rosenblatt, L. M. (1978). *The reader, the text, the poem: The transactional theory of literary work.* Carbondale: Southern Illinois University Press.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research, 64,* Educational Research Association, Chicago, IL. 479–530.

Schuder, R. T. (1993). The genesis of transactional strategies instruction in a reading program for at-risk students. *Elementary School Journal, 94,* 183–200.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* (M. Cole, V. John-Steiner, S. Scriber, & E. Souberman, Eds. and Trans.). Cambridge, MA: Harvard University Press.

Wyatt, D., Pressley, M., El-Dinary, P. B., Stein, S., Evans, P., & Brown, R. (1993). Comprehension strategies, worth and credibility monitoring, and evaluations: Cold and hot cognition when experts read professional articles that are important to them. *Learning and Individual Differences, 5,* 49–72.

(Appendix follows on next page)

Appendix

Summary of "Mushroom in the Rain" Lessons

SAIL Teachers

Teacher 1

The teacher reviewed what expert readers do. She questioned students about the strategies good readers apply when reading. She augmented their responses, explaining some benefits of strategies use. She reviewed with students what they could do when they came to an unknown word (e.g., use picture clues, guess, skip, look back in text). She also focused on verbalizing thinking, summarizing, and visualizing. She asked students to browse through pages and make predictions. A student predicted that the story might be like "The Mitten," a story the group had read earlier in the year. Students discussed possible connections between the two stories. The teacher directed students to verify their predictions as they read and had them visualize a descriptive segment. Students took turns reading. When they finished reading, students either thought aloud spontaneously or were cued to do so by the teacher. Thinking aloud consisted of summarizing content, voicing an opinion, suggesting an interpretation, making or refining predictions, or relating text content to background knowledge or personal experiences. After the reader thought aloud, other students were encouraged to elaborate, persuade, or counter the interpretation.

Students frequently supported their interpretations with background knowledge or text clues. Students continued to discuss similarities and differences between "Mushroom in the Rain" and "The Mitten." For example, they debated whether the mushroom was growing or stretching. Students practiced sequencing by summarizing story content. During discussions, the teacher restated students' responses, clarified confusions, sought elaborations, and garnered opinions from group members. When students faced a word they did not know, they were urged to use one of their "fix-up" strategies. The teacher generally did not ask specific questions about text details. At the end of the lesson, students verified their predictions and fine-tuned their interpretations using text information and background knowledge. Several students admitted they were confused by aspects of the story. When the teacher asked what they could do about this, a student suggested they reread the story. The teacher replied that a good strategy to clarify confusions was rereading. The lesson ended with a student summarizing the story.

Teacher 2

The teacher reviewed what good readers do. Students described the various strategies and evaluated their usefulness. When students talked about visualizing, the teacher explained a personal use of the strategy. The teacher discussed with students the flexible application of a coordinated set of strategies. She encouraged students to use their strategies during story reading. The teacher told students she would focus on visualizing in the lesson. She read the title and first page, modeling her thinking as she visualized text content and made connections between the story and her experiences. She encouraged students to relate the story to their own experiences. Without prompting, a student predicted that the story would be like "The Mitten," a story the class read earlier in the year. The teacher asked the student to support his claim. Students took turns reading aloud. When they came to an unknown word,

they often used strategies without teacher prompting. When they needed help, the teacher cued them to use one of their problem-solving strategies ("fix-it kit"). After reading a page, students would think aloud on their own or be prompted to do so by the teacher. When thinking aloud, students summarized story content, made predictions, or offered interpretations. Other students would then respond to the first student's remarks.

Students continually discussed how "Mushroom in the Rain" was similar to "The Mitten." The group referred to different versions of the story. At one point, a student observed that the animals going under the mushroom were increasing in size. When observations like this one were given by students, the teacher told the group to bear them in mind as they read. Students made and verified predictions frequently and related events to their background knowledge and personal experiences. They elaborated on each other's ideas. During discussions, the teacher did not state her own opinion. Instead, she rephrased students' comments or sought elaboration. When the group thought about what happens to a mushroom in the rain, some students believed the mushroom grew; others countered that the animals stretched it. The teacher allowed students to choose the interpretation they favored. The teacher praised students for their use of strategies, such as making connections between "The Mitten" and "Mushroom in the Rain." She encouraged the group to continue to use strategies in future years because they would help them become better readers.

Comparison Teachers

Teacher 7

The teacher reviewed new words that were presented on cards in the context of sentences. Students were prompted to use the word attack strategies they had been practicing: looking at the first sound, proceeding to the vowel, and then seeing if the word had a suffix. Students took turns reading the story aloud. When students had difficulty, the teacher prompted them to use their word attack strategies and sometimes she gave them the word. After students read, the teacher periodically summarized what had transpired. She drew students' attention to the illustrations. She asked students literal and interpretive comprehension questions about the text, activated their background knowledge, solicited their opinions, and allowed divergence in interpretations (e.g., "Does the ant want to share the mushroom? What does the mushroom remind you of? What do you use in the rain?"). These questions typically did not generate extended discussion. When a student mentioned that the butterfly couldn't fly because his wings were wet, the teacher reminded students of their unit on butterflies. One topic students had been learning about was "persuasion"; the teacher related this topic to the way the animals were persuading the ant to let them under the mushroom. After reading a section, the teacher often asked students what they were thinking. The teacher taught new vocabulary in context, relating word meanings to students' background knowledge.

At one point, the teacher drew a mushroom on the board. She asked students to tell her the order of animals that went under the mushroom. She questioned how all the animals fit under the mushroom. She related this story to other stories students had read. One student said the mushroom grew because of the rain. She

confirmed that mushrooms grow rapidly in the rain. When students faced unfamiliar words, she directed them to apply their word attack strategies and knowledge of phonics (e.g., "Good boy, it's got that double *p* to keep that *o* short . . . ."). After reading one section, she drew students' attention to the quotation marks, colon, commas, and exclamation mark that were on the page. She asked for predictions, without requesting support for students' ideas. Some interpretive discussion occurred around the nature and motives of the fox. When adding the fox to the sequencing on the chalkboard, she said, "When you're making a sequence and you're writing a story or reading it, sometimes it's nice to make an illustration, and then you can add words underneath it to help you organize, get things in, what happened first, second, third, next, and then final." After reading, the teacher frequently drilled students on word skills, using words from the story. Students received a "point" for answering questions correctly. She asked students to find words with suffixes and base words. She frequently provided direct instruction of rules (e.g., making plurals from singular forms; "To keep the *i* short before you add a suffix that begins with a vowel like *-er, -ing, -est*, . . . *-ious*, we have to make sure there's two consonants, to double the letter."). Periodically, she complimented students on their thinking. After reading, students pretended to touch a mushroom. She asked for descriptive words and similes. At the end of the lesson, the teacher told students to visualize to help them remember the ordering of story events. She informed students that they would retell and illustrate the story the next day.

Teacher 10

The teacher stated the title of the story. She asked the students to read the first three pages silently, looking for words they did not know. As students pointed out unfamiliar words, the teacher helped them with word clues. For example, she said that "one of the ways we can find out what a word is sometimes, if we're not too sure of it, is to see if there are little tiny word clues inside of a big word and that will help sound out the word . . . . That's a good word attack skill." The teacher then had a student read the first page. She directed the group to look at the illustration. She told them to notice the size of the mushroom and to watch out for

what happens. There was little discussion during story reading. However, at one point, a student volunteered that the story was like "The Mitten." The teacher did not elaborate on the student's comment except to say "let's see what happens." Toward the end of the story, the teacher asked what happened to the mushroom. One student said the mushroom grew. When the teacher asked why, he answered that it was because, "the water came in the soil and made it grow." At the end of the story, the teacher said that the student "found the secret. That was the secret of how they all fit." Others concurred. One student pointed to the picture of the mushroom getting bigger. The teacher elaborated, "All right, so S found out because he was watching the pictures and getting a clue from the pictures." The group talked a little more about plants needing lots of water to grow. The teacher asked students to tell about any character they liked and what they liked about him. Several students gave opinions.

Discussion then centered on the fox's nature. Students used their prior knowledge to state that the fox was smart. The teacher redirected students to a specific page, asking them to look for a clue. The students recognized that the fox was tricked, and they changed their minds. The group spent much time discussing this episode and looking at the picture. The teacher asked students to fold a piece of paper into four sections. She asked them to draw in order what happened to the mushroom, telling them they could refer to the book for help. She guided them through the activity. The teacher then asked students to suggest alternate endings. Several students responded. She asked students to web the character traits of one of the animals. She told them to "go back into your story and see if there are any story clues . . . and think of some words that would describe that particular character." Students took turns sharing their webs and descriptive words. The teacher asked if they liked the story and whether "it had a nice moral to it. Was it a good lesson about kindness?" Students assented but did not discuss their reasons. She suggested students put new words they learned in their ABC books (i.e., personal word books).