

The Effects of Instruction in an Inference Strategy on the Reading Comprehension Skills of Adolescents with Disabilities Author(s): Nanette Salim Fritschmann, Donald D. Deshler, Jean Bragg Schumaker Source: Learning Disability Quarterly, Vol. 30, No. 4 (Fall, 2007), pp. 245-262 Published by: Council for Learning Disabilities Stable URL: <u>http://www.jstor.org/stable/25474637</u> Accessed: 02/02/2010 09:07

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=cld.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Council for Learning Disabilities is collaborating with JSTOR to digitize, preserve and extend access to *Learning Disability Quarterly*.

THE EFFECTS OF INSTRUCTION IN AN INFERENCE STRATEGY ON THE READING COMPREHENSION SKILLS OF ADOLESCENTS WITH DISABILITIES

Nanette Salim Fritschmann, Donald D. Deshler, and Jean Bragg Schumaker

Abstract. The purpose of this study was to determine the effects of teaching eight secondary students with disabilities, including seven with learning disabilities, a strategy for answering a variety of inferential questions. A multiple-baseline across-subjects design was employed. Outcome measures included scores on researcherdevised comprehension quizzes, a standardized test of reading comprehension, a strategy use test, a strategy knowledge test, and a reading satisfaction measure. Fidelity of implementation, instructional time, and maintenance of skills were also measured. Results suggest that students with disabilities can learn to use a strategy to answer a variety of inferential questions, and mastery of its use can result in improved scores on criterion-based and standardized measures of reading comprehension. In addition, students' satisfaction with their reading improved.

NANETTE SALIM FRITSCHMANN, Ph.D., Lehigh University. DONALD D. DESHLER, Ph.D., University of Kansas. JEAN BRAGG SCHUMAKER, Ph.D., Edge Enterprises, Inc., Lawrence, Kansas.

The current educational climate and its calls for increased skill acquisition and rising performance demands are requiring students to learn higher-order reading skills, like inference skills (e.g., American Institute for Research, 2005; Partnership for 21st Century Skills, 2006). With few exceptions, all secondary students, including students with disabilities (SWD), are required to take rigorous state reading competency exams, most of which involve the use of inference skills. At present, 26 states administer exit exams, and 19 of them withhold diplomas based on poor performance on the exit exams (Center on Education Policy, 2005).

Increased local demands appear to be rising in tandem with the level of reading skills evaluated by national standardized assessment exams. For example, the proposed 2009 National Assessment of Educational Progress (NAEP) reading framework reflects expansion from its current 1992-2007 framework, to include the assessment of broader reading content and deeper cognitive processes (American Institute for Research, 2005). This framework represents a shift from assessing skills at the literal/word level of reading comprehension to assessing higher-order skills within reading comprehension that emphasize "interpreting and integrating" reading matter, the very skills required to make inferences.

This climate of increased reading demands in schools and on tests poses significant challenges for struggling adolescent readers. For students who have a disability, increased demands are especially problematic (Bulgren, Marquis, Deshler, Schumaker, & Lenz, 2006; Schumaker, Deshler, Bui, & Vernon, 2006). Some research has shown that students with learning disabilities (LD) enter seventh grade reading, on average, at the fourth-grade level, and they do not make gains in reading achievement as they progress through the secondary grades (Deshler & Schumaker, 2006; Deshler et al., 2006; Warner, Schumaker, Alley & Deshler, 1980). Further, large proportions of these students are failing their state reading competency exams (Heubert, 2002), as well as tests in their required high school courses (Bulgren, Schumaker, & Deshler, 1988; Hughes, Deshler, Ruhl, & Schumaker, 1993; Wagner et al., 2003).

This is understandable, because, although some of them have acquired some basic decoding skills (Catts, Fey, Tomblin, & Zhang, 2002), they have not learned many of the skills associated with reading comprehension, including inference skills (Gersten, Fuchs, Williams, & Baker, 2001). The combination of more demanding academic requirements and their inadequate reading skills contributes to poor academic outcomes for students with LD (Leach, Scarborough, & Rescorla, 2003; Mastropieri, Scruggs, & Graetz, 2003; Wagner et al., 2003).

Further complicating matters is the fact that higherorder skills, such as those involved in reading comprehension, in general, and making inferences, in particular, can be much more difficult to teach students with LD to a point of proficiency than lower-order processes (Fisher, Schumaker, & Deshler, 2002; Swanson, Hoskyns, & Lee, 1999). Making inferences has been defined as the ability to "construct the text base and the mental models that go beyond the information directly articulated in the text" (Snow, 2002, p. 108). The ability to generate inferences is typically tested by asking questions like, "Why did the boy take action like he did?" or "What is the main message of this passage?"

The answers to such question do not appear directly in the text. Instead, the reader is expected to integrate clues in the text with prior knowledge to create an answer. Thus, although the ability to generate inferences may be critical to text comprehension, the generation of an inference in reading is essentially the result of the *individual* reader's response to the ideas presented in the text; this is somewhat dependent on his/her ability to connect or bridge those ideas with some prior knowledge and with clues provided in text (Pressley, 2000).

To provide a theoretical framework for what happens during inference generation, Kintsch (1998) suggested that during the comprehension process, mental representations are constructed or formed about the information being read in text. When the reader thinks, talks, or writes about these representations, they "undergo integration, which results in a well-structured" (Kintsch, 1998, p. 95) understanding of the text. In the case of drawing inferences from text, Kintsch postulated that various factors (e.g., text features, language skills, and domain knowledge) contribute to comprehension and assist the reader in integrating information into a meaningful structure. When such integration occurs, the reader is able to draw a successful inference related to the text. Also, according to Kintsch, this process may be either automatic (unconscious) or controlled (conscious and strategic).

Research conducted with elementary-level students lends support to Kintsch's theory. For example, some studies have shown that the scores of poor comprehenders on inferential comprehension questions improve when they receive prompts to attend to integrative factors like those highlighted by Kintsch (1998), such as text features and background knowledge (Cain & Oakhill, 1999), and when they are given integrative stimuli (e.g., a descriptive title) along with the passage (Yuill & Joscelyne, 1988). Other research studies in which poor comprehenders have been explicitly taught how to make inferences have focused on teaching students to attend to integrative factors similar to those specified by Kintsch. In each study, students were taught one or some combination of the following skills: activating their background knowledge, making predictions, asking and answering questions, looking for clues in the text, making connections between prior knowledge and information in the passage, and attending to text structure.

Although the results of the studies in this area are somewhat mixed, three studies show positive treatment effects for poor comprehenders (e.g., Dewitz, Carr, & Patberg, 1987; Hansen & Pearson, 1983; Yuill & Jocelyne, 1988). However, none of these studies focused on students with disabilities or on secondary students. The poor comprehenders' average posttest scores on some of the criterion-based tests hovered below the passing range. Furthermore, none of these studies used standardized measures of reading comprehension.

With regard to secondary students, researchers focusing on the use of conscious strategic reading comprehension processes have reported that students with LD can learn to use comprehension strategies that conceivably could contribute to inference generation (Gersten et al., 2001; Swanson et al., 1999). Examples of such comprehension strategies include summarization (Gajria & Salvia, 1992), activating background knowledge and prediction (Afflerbach, 1990), and clarifying (Simmonds, 1992). Other research has shown that secondary students with LD can learn complex reading strategies and that their scores on criterion-based reading comprehension measures can increase as a result of

	Table 1 Student In	formation									
								Comprehe	GRADE nsion Compos	ite Scores	
	Student	Sex	Age Y-M	SES	Race	Disability	Ŋ	Standard Score	Percentile	Grade Equivalent	
	1	W	15-08	Z	Multi-racial	SLD	67	<55	4	2.5	
	2	W	16-05	Ш	Hispanic	LEP/SLD	88	74	4	3.9	
	3	Ţ.	15-01	ĹŦĸ	African-American	MR	40	<55	4	2.3	
	4	Ĭ.	15-10	Z	American- Indian	SLD	76	60	4	3.3	
	5	W	15-02	F	White	SLD	86	66	1	3.3	
	9	۲.	15-05	Z	African-American	SLD	59	<55	4	2.7	
	7	W	15-01	Ĩ.	White	SLD	105	66	1	3.3	
	8	Ĩ.	15-02	N/A	African-American	SLD	85	<55	√1	2.5	
	Means	[15.46				76.0	59.75	1.38	2.9	
	N=neither free n N/A=Not availab	ior reduced-cost lu ile.	inch; F=free lun	ch or reduced-cc	st lunch; SLD= Specific learning o	disability; MR=Mentall	y retarded; LEP:	=Limited English pr	roficiency.		
_											

the strategy instruction (see Schumaker & Deshler, 2006, for a review).

Each of the studies in this area has focused on one reading strategy (e.g., self-questioning, visual imagery) that might be related to inference generation. None has addressed inference generation as an outcome measure or has employed a standardized reading test to measure changes in comprehension. Further, no study to date has investigated teaching a comprehensive package of strategies that might be used for generating several types of inferences.

Thus, the purpose of this study was to develop and test the effects of an instructional program designed to teach an inference strategy to secondary students with disabilities. Specifically, the study was designed to examine the effects of explicit instruction in a multicomponent inference reading comprehension strategy by assessing (a) student knowledge of the strategy, (b) student use of the strategy while reading narrative passages, (c) student ability to answer four types of inferential questions as well as literal comprehension questions, (d) student scores on a standardized measure of reading comprehension, (e) student reading and strategy satisfaction, and (f) required instructional time for students with disabilities in a secondary setting.

METHOD

Participants

Participants were eight ninth-graders with disabilities whose parents had given consent for their participation. Students were enrolled in learning-supported English/language arts classes. A standardized reading test, the Group Reading Assessment and Diagnostic Evaluation (GRADE) (Williams, 2001), indicated that the students' reading scores fell at least five grade levels below their current grade placement. Each student had been designated as a student with a disability and had been placed in a resource program for a minimum of 180 minutes per day on the recommendation of a multidisciplinary special services team with documented parent/guardian approval (see Table 1 for demographic and test data on the students). Of the eight participants, seven had been classified by their school district as having a learning disability; one had been classified as having mental retardation. The participating district followed the IQ-achievement discrepancy model for identification of learning disabilities, requiring a minimum discrepancy of 18 points between a student's IQ and achievement scores (J. Harrington, personal communication, July 21, 2007). Further, the state within which the participating district is located is one of five states in this country that rely on the professional judgment of a team in the determination process (Reschly, Hosp, & Schmied, 2005).

Setting

The school district is located in an urban midwestern community with a population of approximately 124,000. Instruction and testing took place in a classroom in the high school and was conducted with two groups of four students each during different class periods.

The Inference Strategy

The Inference Strategy is a reading comprehension strategy designed to help students create meaning from clues provided in text and respond to a variety of inference questions. Standardized reading tests were analyzed to determine the types of inference questions that students were expected to be able to answer. Four main types emerged: purpose, main idea/summarizing, predicting, and clarifying (see Figure 1 for examples).

The Inference Strategy taught to the students in this study consists of five steps. During Step 1, "Interact with the passage and the questions," students first preview the passage, paying particular attention to the title and the length of the passage. Then they read the questions and mentally identify two main *categories* of questions: factual questions and think-and-seek (inferential) questions. Next, the students further classify the thinkand-seek questions into four types: purpose, main idea/summarization, prediction, and clarification questions. During the second step, "Note what you know," students activate any background knowledge or experiences they may have related to the topic and questions, underline any key words in the questions that indicate what information to look for in the passage, and note code letters next to each question to indicate the category of the question and the question type. During the third step, "Find the clues," students carefully read the passage and find and underline clues that are directly related to key words in the questions. They then create tentative answers to the questions mentally. The fourth step, "Explore more details," prompts students to look for any additional clues in the passage that support the tentative answers they have selected. The final step, "Return to the question," calls for the students to go back to each question and make sure that an answer has been selected and marked.

Thus, these steps were designed to cue students to attend to their prior knowledge, to attend to the type of inference they were being asked to make, to attend to key words in the questions that would help them search for clues in the text, to search for those clues, and to problem solve once they had gathered the clues to make inferences about the information in the passage. The use of the generic strategy steps was constant regardless of the type of question the student was addressing. However, students looked for different types of clues,



depending on the type of think-and-seek question. (See the Procedures section for a description.)

The mnemonic device "INFER" was created from the first letters of the steps to help students to remember the steps' names and their order in the strategy. Initially, the strategy steps are to be used in the "I" to "R" order; however, thereafter, the steps may be used in a recursive and flexible manner to allow students to cycle back to any step if necessary.

Instruments and Measures

Fidelity checklists. A checklist was used to measure teacher adherence to the instructional sequence for the lessons. It listed several teacher behaviors: provide an advance organizer, discuss the purpose of the lesson and provide rationales for the lesson, state expectations for

student behavior, describe a step of the strategy or how to use the strategy, model the strategy, provide practice opportunities with feedback, and provide a post organizer.

The number of these behaviors varied for each lesson's checklist (from 7 to 11), depending on the content of the lesson. For example, if one type of inferential question was to be introduced in a given lesson, there were seven items on the checklist, corresponding to those listed above. If two question types were to be introduced in a given lesson, then the strategy was described and modeled twice, once for each question type, and nine items were listed on the checklist.

The delivery of each of the instructional lessons was recorded using a tape-recorder and audiotapes. The checklist was filled out by a scorer while listening to **Strategy use test.** To obtain repeated measures of students' use of the Inference Strategy, a pool of 30 ninth-grade-level narrative passages from the *Jamestown Readers – Timed Readings in Literature Series* (Spargo, 1989) was created; passages were randomly selected and sequenced for each student from this pool. This level of passages (i.e., ninth grade) was chosen to provide information on performance at the students' current grade level because this is the level at which they are expected to perform in general education classes. Narrative passages were selected to control for the type of passage as well as to fulfill requirements for the English class in which the students were enrolled.

For each probe test, the student was asked to read three of the 400-word passages silently and use the Inference Strategy in relation to the passage and the five factual and inferential questions that followed the passage (see the next section for more about these questions). To measure use of the strategy, students were awarded one point for each of the following strategic behaviors: underlining key words in a question, underlining clue words in the passage related to the question, recording a code letter identifying the *category* of question (i.e., factual versus think-and-seek), and recording a code letter identifying the *type* of think-and-seek question (i.e., purpose, main idea/summarization, prediction, or clarification).

Scorers used an answer key for awarding a total of four points per think-and-seek question and three points per factual question. (Students classified think-and-seek questions by type. They did not have to classify the factual questions by type, so they could not earn a fourth point for these questions.) A total of 19 points were available per passage for the strategy use score or 57 points per probe test (since three passages were read per probe test).

Criterion-based comprehension test. As mentioned, each passage the students read was followed by five multiple-choice questions written by a researcher specifically for a given passage. The five questions consisted of one factual question and one question for each of the four types of think-and-seek questions: purpose, main idea/summarization, predicting, and clarifying. Student answers to these questions were the criterion-based comprehension measure used in this study. Students were awarded 1 point per correct answer (based on an answer key) for a total of 5 points per passage and a total

of 15 points per probe test (since there were three passages per probe test).

Strategy knowledge test. The Strategy Knowledge Test was used to measure the students' knowledge of the steps of the Inference Strategy as a pretest and posttest measure. It included five short-answer questions, which required students to list and explain the strategy steps and their uses. An answer key specified the parameters for correct answers. For example, Question #1 was, "What is the first step a reader takes when using the INFER Strategy?" (The correct answer was "Interact with the questions and the passage.") For each answer, a student could earn either zero points (no information), 5 points (partial information [e.g., the student wrote "Interact with the questions"]), or 10 points (completely correct information [e.g., the student wrote the whole name of the step]). A total of 50 points was available on the test.

Standardized reading test. Two subtests of the GRADE (Williams, 2001) were administered to the students prior to and after instruction: Sentence Completion and Passage Comprehension. The combined scores from these subtests comprise the GRADE comprehension composite score, which was the score used in this study. Forms A and B of the GRADE were used for pre- and posttesting, respectively. Reliability coefficients for the alternate forms and test-retest are in the .90 range. In addition, correlations between scores on the GRADE and on the Gates-MacGinitie Reading Test "range from .81 to .94, with half of the coefficients being .89 or higher" (Williams, 2001, p. 85).

Student satisfaction survey. The Student Satisfaction Survey consisted of 10 questions, each formatted using a 7-point Likert-type scale, ranging from "1" ("Totally wrong") to "7" ("Totally right"). This instrument was administered to obtain a pretest and a posttest measure. On survey items, the students were asked to rate their attitudes toward reading in school and learning and toward using the Inference Strategy for the purpose of reading passages and responding to comprehension questions. Example survey items are "I am a good reader in school," "I know what steps I can take to make meaning from what I read," and "I feel that I can use the Inference Strategy to help me understand what I read in class."

Time required for instruction. Instructional time was recorded in a journal kept by the researcher. Start and stop times, including hours and minutes, were recorded, along with dates of instruction. Teacher time began when the teacher started (or restarted) the lesson with the students. It ended when an interruption occurred (e.g., a phone call), or when the students began practicing the strategy. Student time began when the teacher started (or restarted) the lesson with the students. It

ended when an interruption occurred, or when the students stopped practicing and handed in their work. Thus, several start and stop times were potentially recorded for each lesson.

Reliability. For the fidelity checklists, two scorers independently scored 40% of the lessons, and their scores were compared item by item. An agreement was tallied if both scorers had recorded the same score on an item. The percentage of agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100. The scorers agreed on 38 out of 39 possible agreements for a total percentage of agreement of 97% (range = 91% to 100%).

Similar reliability checks were completed for the Strategy Use Test, the criterion-based comprehension tests, and the Strategy Knowledge Test for 40% of the tests taken during baseline and post-intervention conditions. Scorers were blind as to which tests were taken during baseline and after instruction, respectively. With regard to the Strategy Use Test, the scorers agreed on 949 out of 969 possible agreements, for a total percentage of agreement of 98% (range = 93% to 100.0%). On the criterion-based comprehension tests, the scorers agreed on 100% of the 255 items scored. Finally, on the Strategy Knowledge Test, the scorers agreed on 28 items out of a possible 30, for a total percentage of agreement of 93% (range = 80% to 100%).

Reliability for teacher and student instructional time was determined for three of the instructional sessions during which two scorers recorded the times separately. Agreement was registered each time the two scorers' times agreed on a start or stop time to the minute. The total percentage of agreement on teacher time was 100%, and on student time it was 93%. The scorers' times agreed on 10 out of 10 possible agreements for teacher times and on 15 out of 16 possible agreements for student times.

Procedures

Instructor. The instruction was completed by the first author, who is a certified special education teacher with five years' experience teaching students with disabilities. She also is a certified strategic instruction model professional developer with the University of Kansas Center for Research on Learning.

Instructional materials. An instructional protocol (Fritschmann et al., in prep.) was written to ensure that instruction was standardized across the two classes. It was comprised of scripted step-by-step instructions for each lesson plus visual devices to be used during the instruction.

Two types of practice materials were constructed. For the first type, nine short, one-paragraph passages were written by a researcher followed by two to four multiple-choice questions corresponding to the type of question(s) covered in each of the lessons. These short passages were used for practice activities immediately after students had been introduced to how to use the strategy for a new question type.

For the second type of practice material, narrative passages from the fourth-, sixth- and eighth-grade levels of the *Jamestown Readers – Timed Readings in Literature series* (Spargo, 1989) were selected, so students could start practicing the strategy with relatively easy passages (the fourth-grade passages). Gradually, across practice attempts, they progressed to more difficult passages (sixth-grade passages), and then to passages written one grade level below their current grade (eighth-grade passages). The length of passages ranged from 200 to 400 words, increasing with the reading level of the passage. For each passage, five multiple-choice questions (one factual and four inferential questions, each corresponding to one of the four types of think-and-seek questions) were written by a researcher.

These long passages and accompanying questions were used after all the question types had been introduced to the students so that they could practice using the strategy with a mixture of questions. Student performance on these activities was scored using an answer key and the guidelines described under the Measures section for the Strategy Use Test and the criterion-based comprehension test.

Pretest procedures. Students completed a Reading Satisfaction Survey and a minimum of three probe tests containing a total of nine Strategy Use Tests and nine corresponding five-item criterion-based comprehension tests over a one-week period. For each probe test, a Strategy Use Score and a Comprehension Test Score were determined. When the third baseline probe was completed, four students whose baselines were stable (hereafter referred to as Cohort A) began instruction in the Inference Strategy. Once the Cohort A students showed an increase in their use of the strategy, the four other students (Cohort B) completed at least one additional baseline probe until their baselines were stable. Then they began the instruction.

Intervention procedures. The students received instruction in the Inference Strategy in sessions ranging in length from 60 to 75 minutes, depending on the school schedule. Instruction was based on a validated instructional methodology for teaching learning strategies to students with disabilities (Schumaker & Deshler, 2006).

In the first instructional session, the students were asked to make a commitment to actively learn and use the Inference Strategy. Also, in the first class session, the instructor explained and described in detail the steps of the Inference Strategy. This was followed by instruction that included how to identify and code the two main categories of questions, (a) factual questions (i.e., questions with answers that are "right there" in the passage) and (b) "think-and-seek" questions (i.e., questions that require the reader to really think about and seek out clues so that they can infer an answer). Additionally, the students were instructed on how to identify and mark each type of think-and-seek question with code letters.

The students were then taught through description and demonstration how to look for key words embedded in the question to code and respond to factual questions. At the end of the first session, the students completed practice activities for coding question categories and answering factual questions for short passages, and they received feedback on their work. If they met mastery (i.e., earned 80% or more of the points), they proceeded to the next instructional session. If they did not reach mastery on the practice activity, they completed additional practice activities until they met mastery.

During the second instructional session, the students were provided detailed information on purpose questions and main idea/summarizing questions, and they were taught how to identify and mark those questions with code letters. They were also taught how to look for key words embedded in the questions and clues imbedded in the text and how to correctly respond to these types of questions. For example, for purpose questions, students were taught that authors may have three main purposes as they write a passage: to entertain, to inform, and to persuade. Each type was defined, and students were taught to look for the key words "author's purpose" or "author's reason" in the question. Then as they read the passage, they were taught to ask themselves "Why do I think the author wrote this?" and attend to certain types of clues that would help them determine the author's purpose. For example, for informative passages, they were taught to look for large amounts of facts and details like they might see in a textbook. For entertaining passages, they were taught to look for sections that made them happy or fearful. After instruction, discussion, and modeling, students were provided with short practice passages followed by a mixture of factual, purpose, and main idea questions, and were given feedback on their efforts. Again, mastery was required before students proceeded.

The third and fourth instructional sessions focused on instruction of and practice with predicting and clarifying questions, respectively. As in the previous sessions, the researcher explained in detail how to identify the new type of question and search for clues in the passage to support a correct response to that type of question. This process was modeled for the students prior to the practice activities with the short passages. The instructor provided group and individual feedback during and after each practice activity, and mastery was required.

In subsequent sessions, students were provided practice activities with the longer fourth-grade-level reading passages, in which they were required to use all the steps of the strategy in response to all question types. If the student earned a score at or above the mastery levels (80% on the Comprehension Test, 70% on the Strategy Use Test), they were moved up to the next reading level (sixth-grade level, then eighth-grade level). They continued to practice and receive individual feedback until they reached mastery on a passage written at the eighth-grade level. (See Fritschmann, 2006, for more details on the instruction.)

Posttest procedures. Upon completion of instruction in the Inference Strategy and reaching the mastery criteria on one eighth-grade practice activity, students took a posttest containing three 400-word, ninth-grade passages. After reading each passage and using the strategy, students were asked to complete a fivequestion criterion-based comprehension test. Each student completed a minimum of one such posttest (containing a total of 3 passages and 15 questions).

Following these tests, the students were administered, as a group, the two selected subtests included in Form B of the GRADE (Williams, 2001). They were also administered the Student Satisfaction Survey and Strategy Knowledge Test in a group setting. The students completed each instrument independently, and there was no time limit for test completion. This posttesting occurred during the last weeks of the school year.

Maintenance procedures. Two maintenance tests were administered during the next academic year. The first was administered eight months after the posttests to six of the eight original subjects who were present on the day the test was administered. The students did not receive a review of the Inference Strategy prior to taking the first maintenance test, nor had the strategy been reviewed with them since they had completed the study eight months earlier. There was no time limit on the test.

A second maintenance test was administered to four students who were still enrolled in the school 12 months after the posttests. In this case, the students took part in a 45-minute review session where they practiced using the Inference Strategy the day before taking the maintenance test.

Research Design

A multiple-probe-across-subjects design (Horner & Baer, 1978) was employed to determine the effects of instruction on students' strategy use and reading com-

prehension performance. All students were given at least three reading probes before instruction began. The students who served as the second tier in the design (Students 2, 4, 6, and 8) had at least four probes before instruction began. Their instruction commenced only after the scores of students in the first tier of the design had improved. All students' baselines were stable before they began the instructional program. Additionally, a pretest-posttest analysis was employed to compare the standardized reading test scores earned before and after instruction on the GRADE.

RESULTS

Fidelity of Implementation Results

The instructor emitted 86 of 88 listed behaviors on the fidelity checklists for a total percentage of implementation of 98%.

Strategy Use Test and Criterion-Based Comprehension Test Results

Figures 2-4 display the percentage of points earned by each of the eight participants on each Strategy Use Test (diamond symbols) as well as the percentage







of comprehension questions answered correctly (square symbols) on each criterion-based comprehension test. (Each symbol represents performance on three passages and three sets of questions.) The grade level at which each reading passage was written is shown along the x-axis of each graph. Student 3's data were graphed with those of Students 4 and 5 because his graph partner moved from the school and could not continue with the study. Likewise, Student 6's data were graphed with those of Students 7 and 8 because his graph partner was excluded midway from the study due to a high absentee rate. The figures show that the percentage of comprehension questions answered correctly increased with the onset of instruction and maintained a positive trend through the remainder of instruction and posttesting. During baseline, the students answered an average of 31.74% of the comprehension questions correctly. During instruction, they answered an average of 77.39% of the questions correctly; during the posttest condition, they answered an average of 82% of the questions correctly. Similar results were evidenced with regard to the Strategy Use Test. During baseline, the students earned an average of 0% of the points on the Strategy Use Test. During instruction, they earned an average score of 66.39%; during the posttest condition, they earned an average score of 81.94% on the Strategy Use Test.

Table 2 displays the results from the maintenance tests. On the left-hand side are listed the comprehension (M = 41%, SD = 11.73) and strategy use percentage scores (M = 19%, SD = 14.19) earned by the six participants who took a maintenance test eight months after instruction was completed without participating in a review (Maintenance Test 1). On the right-hand side are the scores earned by the four participants who took the test 12 months after instruction was completed and after a brief review (Maintenance Test 2). The scores earned by the latter group of students suggest that with a brief review and practice, students were able to apply the steps of the strategy and answer questions correctly

at levels that were higher than during baseline and that approximated their posttest performance more closely than when they took the test without a review.

A Friedman Test was conducted to evaluate differences between the median for the Strategy Use Test scores during baseline (median = 0%), posttest (median = 82%), and maintenance without review (median = 19 %) conditions. (Data from the maintenance withreview test were not included because of the low number of subjects.) Significant differences were found, χ^2 , (2,*N* = 6) = 11.565, *p* < .01, and the Kendall coefficient of concordance (effect size index) of .964 indicated strong differences among the three median scores.

Follow-up pairwise comparisons were conducted using a Wilcoxon Signed-Ranks Test (Siegel & Castellan, 1998); the LSD procedure was used to control Type I errors across these comparisons at the .05 level. The

		<u> </u>		
Student	Maintenan Comprehension Score Without Review ^a	ce Test 1 Strategy-Use Score Without Review ^a	Maintenanc Comprehension Score With Review ^b	e Test 2 Strategy-Use Score With Review ^b
1	40	0	60	79
2	n/a	n/a	67	84
3	26	15	53	77
4	60	22	n/a	n/a
5	47	40	n/a	n/a
6	33	29	n/a	n/a
7	40	10	n/a	n/a
8	n/a	n/a	80	88
Mean (SD)	41 (11.3)	19 (14.19)	65 (11.52)	82 (4.97)

Student	Pretest	Posttest	
1	0	100	
2	0	96	
3	0	80	
4	0	100	
5	0	96	
6	0	92	
7	0	80	
8	0	90	
Mean (SD) 0	91.75* (8.031)	

median score for the Strategy Use posttests was significantly higher than the median score for the baseline tests, p = .012, and it was also significantly higher than the median score for the maintenance without-review tests, p = .028. The median score for the maintenance without-review test was significantly higher than the median baseline score, p = .043.

A Friedman Test was also conducted to evaluate the differences in the medians for the percentage of comprehension questions answered correctly during the baseline (median = 33.51 %), posttest (median = 78.42%), and maintenance-without-review conditions (median = 41.00%). Significant differences were found, $\chi^2(2,N=7) = 12.00$, p < .01, and the Kendall coefficient of concordance (effect size index) of 1.00 indicated strong differences among the three median scores. Follow-up pairwise comparisons were conducted using a Wilcoxon Signed-Ranks test (Siegel & Castellan, 1998); the LSD procedure was used to control for Type I errors across these comparisons at the .05 level. The median score on the posttest comprehension tests was signifi-

cantly higher than the median score for the baseline comprehension tests, p = .012, and it was also significantly higher than the median score for the maintenance test, p = .028. The median maintenance score was significantly higher than the median baseline score, p = .027.

Strategy Knowledge Test Results

Table 3 lists the percentage scores earned by the eight students on the Strategy Knowledge pretest and posttest. As illustrated, each student earned a score of 0% on the pretest (M = 0.00%, SD = 0.00). Percentage scores on the posttest ranged from 80% to 100% (M = 91.75%, SD = 8.031). The Wilcoxon Signed-Ranks Test (Siegel & Castellan, 1988) was used to determine whether there were significant gains from pretest to posttest on this measure; the LSD procedure was used to control Type I errors across these comparisons at the .05 level. A significant difference was revealed between the pretest and posttest scores, z = -2.530, p = .00. The effect size for this gain was r = 0.99, representing a large gain according to Cohen (1988).

Standardized Reading Test Results

Table 4 lists the mean standard scores and grade-level scores earned by students on the pretest and posttest on the GRADE. The Wilcoxon Signed-Ranks Test revealed a significant difference between the pretest and posttest standard scores, z = -2.521, p = .012. The effect size for this gain was r = 0.91, a large gain according to Cohen (1988). This gain represents an average increase of 2.82 grade levels (range = 1.4 to 3.6 grade levels) in reading comprehension.

Satisfaction Results

The mean pretest and posttest reading satisfaction ratings are reported in Table 5. As illustrated, significant differences were found between the pretest and posttest mean scores for the 7-point Reading Satisfaction Questionnaire using a Wilcoxon Test, z = -2.524, p = 0.12. The effect size for this difference was r = 0.95, a large gain according to Cohen (1988).

Table 6 displays the mean pretest and posttest ratings and the standard deviations for each item on the questionnaire. The items that received the highest posttest ratings related to enjoying the exercises and passages while learning the strategy (M = 6.62) and using the strategy to aid understanding in class (M = 6.37).

Time Required for Instruction

Instructor time. The total amount of instructor time required to deliver the initial Inference Strategy instruction ranged from 280 min. to 350 min. (M = 300 min., or approximately 5 hours).

Student time. The total amount of student time included time spent working with the instructor (see the time reported above) and working independently on practice activities with reading passages. It ranged from 770 min. to 1040 min. (M = 905 min., approximately 15 hours).

Table 4 The GRADE Comprehension Standard Scores and Grade-Level Equivalents GRADE GRADE **Grade-Level** Grade-Level Student Pretest SS Posttest SS **Equivalent Pre Equivalent Post** 55 88 2.5 5.6 1 2 74 88 3.9 5.6 3 87 2.3 5.3 55 72 94 3.8 7.2 4 5 66 83 3.3 4.7 6 55 91 2.7 6.3 7 58 89 2.9 5.9 2.5 8 55 89 5.9 2.99 (0.52) Mean (SD) 61.25 (8.17) 88.63* (3.16) 5.81 (0.73) *Significant difference found

· · · · · · · · · · · · · · · · · · ·			
Student	Pretest	Posttest	
1	1.1	5.8	
2	1.6	4.9	
3	1.5	5.1	
4	2.3	5.5	
5	1.5	4.4	
6	1.8	6.4	
7	1.3	6.5	
8	1.3	6.6	
Mean (SD)	1.55 (3.75)	5.65* (.815)	

DISCUSSION

Several conclusions may be drawn from the results of this study. First, instruction in the Inference Strategy increased students' use of strategic skills to a mastery level. Second, the instruction produced a positive change in the students' ability to respond to inferencetype questions on criterion-based tests. Increases in strategy use and ability to answer inferential questions were found only after students participated in the Inference Strategy instruction as demonstrated through the multiple-baseline design. Third, students' posttest scores on the GRADE were significantly higher than their pretest scores following instruction in the Inference Strategy. Moreover, the GRADE scores indicated that, on average, the students made a 2.8 gradelevel gain in reading comprehension within 15 hours of instruction. The effect sizes related to the gains associated with all three of these major outcome measures were large. During the instruction, the students gradually worked up to reading passages written at their grade level and responded to associated inference comprehension questions at mastery levels. They also earned scores above 90% on a test of their strategy knowledge at the end of the study.

In addition, the results of the Student Satisfaction Questionnaire suggest that students who participated in the Inference Strategy instruction were more satisfied at the end of the study than at the beginning of the study with how they felt about reading and different comprehension processes. On the one question related to their reaction to the instruction, the students indicated that they enjoyed the instruction (M = 6.62 on a 7-point scale). Finally, instruction of the Inference Strategy required five hours of initial instructional delivery and then another 10 hours of supervising practice activities and providing feedback.

Table 6

Mean Pretest and Posttest Ratings, Standard Deviations, and Ranges for All Students on Each Item on the Reading and Strategy Satisfaction Questionnaire

Fill in the box that best describes how you feel		Pretest	Posttest
Reading in school is boring.	Mean	6.12	4.375
	SD	0.83	2.39
	Range	5-7	1-7
I am a good reader in school.	Mean	2.00	3.63*
	SD	1.07	2.64
	Range	1-4	1-6
Reading assignments in school confuse me.	Mean	6.37	6.00
	SD	1.06	1.19
	Range	4-7	1-7
I know what an inference in reading is.	Mean	1.37	4.75*
	SD	0.52	2.25
	Range	1-2	1-7
Coming to a conclusion from what I read in class	Mean	6.50	5.37
can be difficult for me.	SD	0.76	1.60
	Range	5-7	3-7
I know what steps I can take to make meaning	Mean	2.06	4.56*
from what I read.	SD	0.64	1.92
	Range	1-3	2-7
I enjoy reading in class.	Mean	2.25	3.75*
	SD	1.28	2.12
	Range	1-4	1-6
I felt comfortable using the Inference Strategy	Mean	1.00	5.87*
while I read in class.	SD	0	1.55
	Range	1	3-7
I feel that I can use the Inference Strategy to help	Mean	1.00	6.37*
me understand what I read in class.	SD	0	0.52
	Range	1	6-7
I enjoyed the exercises and passages that were	Mean	1.00	6.62*
used while learning the Inference Strategy.	SD	0	0.52
	Range	1	6-7
* Represents a mean gain of at least one point.			

Several limitations apply to this study. First, although the sample size is adequate for the experimental design used, it was small. Expanding the research design to include more classrooms would increase the sample size and, therefore, eliminate the problem of committing a Type II error and making generalizations that are not based on a sufficiently large sample of students.

Another possible limitation relates to the characteristics of the participants. They represent students with severe reading comprehension deficits, beyond those typically evidenced in the population of students with learning disabilities at large. How much instructional time would be required and the kinds of gains that might be made by students exhibiting lesser deficits still need to be determined. Another limitation is that test scores on the students' decoding skills were not available. Thus, it is unknown whether they had higher levels of decoding skills than comprehension skills, which enabled them to tackle the higher reading-level passages.

An additional concern is that the instruction was provided by a researcher. Whether other teachers can produce the same types of reading gains is unknown. Further, expository passages were not included in the instructional materials or the criterion-based tests. Although they were included on the GRADE, how the students performed on those particular passages is not known. Whether the strategy can be used successfully with expository passages remains to be determined. This is an important consideration because expository passages represent a considerable portion of the materials used in secondary content-area classes and are included within state and national assessments. Another concern is that the results were not disaggregated according to the type of question on the tests. Thus, at this time, whether the students responded more successfully to some types of questions than others is not known.

A final limitation relates to the maintenance data collected eight months after the initial intervention. While these data were collected on a relatively small subset of the entire sample, some trends require an explanation. Specifically, performances on both strategy use and reading comprehension probes were relatively low. Educational research has shown that the maintenance of a targeted behavior as a result of an intervention has largely been an unrealized goal of behavioral interventions (McConnachi & Carr, 1997). While features of the intervention itself need to be considered as possible reasons for the low performance, other factors may be responsible. In this study, the students received no additional instruction in the use of the strategy nor prompts to use it after the study was completed. These maintenance data seem to underscore the importance of continued practice and periodic maintenance probes if students are expected to continue to use a given strategy.

On the one hand, these data indicate that comprehension strategies learned by students with disabilities will not continue to be an active part of their repertoire of reading behaviors unless students are given opportunities to practice the strategies over a sustained period of time. On the other hand, these data are encouraging, because they clearly demonstrate that proficiency with the targeted behavior can be quickly regained when a relatively modest review procedure is used. Clearly, the instructional dynamics surrounding maintenance and sustained use of complex strategies such as the Inference Strategy require considerably more study.

Research is needed to determine the effects of Inference Strategy instruction with larger, more diverse, and more carefully described groups of students, and with teachers, as opposed to researchers, providing the instruction. Maintenance procedures need to be explored in more depth and detail. In addition, further analysis of the data focusing on student performance on each of the question types may provide helpful information. Last, research is needed that focuses on use of the strategy with expository as well as narrative passages.

Implications for Practice

The instruction described here for the Inference Strategy has the potential of impacting education more broadly in the following ways. First, the instruction can potentially serve as a vehicle through which small groups of students with disabilities can be taught the inference skills required on today's standardized reading tests. As a result, schools will be better equipped to meet increasing requirements and students' needs. Second, teachers may be able to combine this type of instruction with other forms of explicit reading comprehension strategy instruction (e.g., see Schumaker & Deshler, 2006, for a review) because of similar instructional design features. Third, the Inference Strategy instruction might be used with student populations representing low socio-economic and ethnically diverse groups, as well as those who have severe reading deficits.

REFERENCES

- Afflerbach, P. P. (1990). The influence of prior knowledge on expert readers' main idea construction strategies. *Reading Research Quarterly*, 25(1), 31-46.
- American Institute for Research. (2005, Spring). *Reading framework* for the 2009 National Assessment of Educational Progress, prepublication edition. Washington, DC: Author. Retrieved March 3, 2006, from http://www.nagb.org/pubs/pubs.html
- Bulgren, J. A., Marquis, J. G., Deshler, D. D., Schumaker, J. B., & Lenz, B. K. (2006). Instructional context of inclusive secondary

general education classes: Teachers' instructional roles and practices, curricular demands, and research-based practices and standards. *Learning Disabilities: A Contemporary Journal*, 4(1), 39-66.

- Bulgren, J. A., Schumaker, J. B., & Deshler, D. D. (1988). Effectiveness of a concept teaching routine in enhancing the performance of LD students in secondary-level mainstream classes. *Learning Disability Quarterly*, 11(1), 3-17.
- Cain, K., & Oakhill, J. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing: An Interdisciplinary Journal*, 11(5), 489-503.
- Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research, 45,* 1142-1157.
- Center on Educational Policy. (2005). *High school exit exams. Policy brief 1, December.* Retrieved April 1, 2006, from http://www.cep-dc.org/highschoolexit
- Deshler, D. D., & Schumaker, J. B. (2006). *Teaching adolescents with disabilities: Accessing the general education curriculum*. New York: Corwin Press.
- Dewitz, P., Carr, E., & Patberg, J. (1987). Effects of inference training on comprehension and comprehension monitoring. *Reading Research Quarterly*, 22(1), 99-121.
- Fisher, J. B., Schumaker, J. B., & Deshler, D. D. (2002). Improving the reading comprehension of at-risk adolescents. In C. C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 351-364), New York: Guilford.
- Fritschmann, N. S. (2006). The effects of instruction in an inference strategy on the reading comprehension skills of adolescents with learning disabilities. Unpublished doctoral dissertation, University of Kansas, Lawrence.
- Fritschmann, N. S., Schumaker, J. B., & Deshler, D. D. (in preparation). *INFER: The inference strategy instructors manual.*
- Gajria, M., & Salvia, J. (1992). The effects of summarization instructions on text comprehension of students. *Exceptional Children*, 58(6), 508-520.
- Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research*, 71(2), 279-320.
- Hansen, J., & Pearson, P. D. (1983). An instructional study: Improving the inferential comprehension of good and poor fourth-grade readers. *Journal of Educational Psychology*, 75(6), 821-829.
- Heubert, J. P. (2002). *Disability, race and high-stakes testing.* Wakefield, MA: National Center on Accessing the General Curriculum. Retrieved March 30, 2006, from http://www. cast.org/publications/ncac/ncac_disability.html
- Horner, R. D., & Baer, D. M. (1978). Multiple-probe technique: A variation of the multiple baseline. *Journal of Applied Behavior Analysis*, 11(1), 189-196.
- Hughes, C. A., Deshler, D. D., Ruhl, K. L., & Schumaker, J. B. (1993). Test-taking strategy instruction for adolescents with behavior disorders. *Journal of Emotional and Behavioral Disorders*, 1(3), 188-199.
- Kintsch, W. (1998). Comprehension: A paradigm for cognition. New York: Cambridge University Press.
- Leach, J. M., Scarborough, H. S., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology*, 95(2), 211-224.
- Mastropieri, M. A., Scruggs, T. E., & Graetz, J. E. (2003). Reading comprehension instruction for secondary students: Challenges

for struggling students and teachers. Learning Disability Quarterly, 26(2), 103-116.

- McConnachi, G., & Carr, E. G. (1997). The effects of child behavior problems on maintenance of intervention fidelity. *Behavior Modification*, 97(2), 123-158.
- Partnership for 21st Century Skills. (2006). *Learning for the 21st century. A report and mile guide for 21st century skills*. Retrieved March 29, 2006, from http://www.21stcentury skills.org.
- Pressley, M. (2000).What should comprehension instruction be the instruction of? In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research, Volume II* (pp. 545-562). Mahwah, NJ: LEA.
- Reschly, D. J., Hosp, J. L., & Schmied, C. M. (2005). And miles to go ...: State SLD requirements and authoritative recommendations. Retrieved July 31, 2007, from www.nrcld.org/research/states/ index.shtml
- Schumaker, J. B., & Deshler, D. D. (2006). Teaching adolescents to be strategic learners. In D. D. Deshler & J. B. Schumaker (Eds.), *Teaching adolescents with disabilities: Accessing the general education curriculum* (pp. 121-156). New York: Corwin Press.
- Schumaker, J. B., Deshler, D. D., Bui, Y., & Vernon, S. (2006). High schools and adolescents with disabilities: Challenges at every turn. In D. D. Deshler & J. B. Schumaker (Eds.), *Teaching adolescents with disabilities: Accessing the general education curriculum* (pp. 1-34). New York: Corwin Press.
- Siegel, S. & Castellan, J. N. (1988). Nonparametric statistics for the behavioral sciences (2nd ed.). Boston: McGraw-Hill.
- Simmonds, E.P.M. (1992). The effects of teacher training and implementation of two methods of improving the comprehension skills of students with learning disabilities. *Learning Disabilities Research and Practice*, 7(4), 194-198.
- Snow, C. E. (2002). Reading for understanding: Toward a research and development program in reading comprehension. Santa Monica, CA: Rand.
- Spargo, E. (1998). Timed readings in literature Jamestown Series. Providence, RI: Glencoe/McGraw-Hill of McGraw-Hill Education.
- Swanson, H. L., Hoskyn, M., & Lee, C. (1999). Interventions for students with learning disabilities: A meta-analysis of treatment outcomes. New York: Guilford.
- Wagner, M., Marder, C., Blackorby, J., Camento, R., Newman, L., Levine, P., & Davis-Mercier, E. (with Chorost, M., Garza, N., Guzman, A., & Sumi, C.). (2003). The achievements of youth with disabilities during secondary school. A report from the National Longitudinal Transition Study-2 (NLTS2). Menlo Park, CA: SRI International. Retrieved April 3, 2006, from http://www. nlts2.org
- Warner, M. M., Schumaker, J. B., Alley, G. R., & Deshler, D. D. (1980). Learning disabled adolescents in the public schools: Are they different from other low achievers? *Exceptional Education Quarterly*, 1(2), 27-35.
- Williams, K. T. (2001). *GRADE: Group Reading Assessment and Diagnostic Evaluation*. Shoreview, MN: Pearson AGS Globe.
- Yuill, N., & Joscelyne, T. (1988). Effect of organizational cues and strategies on good and poor comprehenders' story understanding. *Journal of Educational Psychology*, 80(2), 152-158.

Please address correspondence to: Nanette S. Fritschmann, School of Education, Lehigh University, Bethlehem, PA 18015; nfritsch@lehigh.edu