Journal of Educational Psychology
2000, Vol. 92, No. 3, 426–441

0022-0663/00/$5.00 DOI: 10.1037//0022-0663.92.3.426

The Use and Effectiveness of Analogical Instruction in Diverse Secondary Content Classrooms

Janis A. Bulgren, Donald D. Deshler, Jean B. Schumaker, and B. Keith Lenz
University of Kansas

The purpose of these studies was to explore the use of analogies while teaching important concepts in secondary content classrooms containing students of diverse abilities. Various research methodologies were used to determine the effects of an analogically based routine on student learning and the effects of training on teacher use of the routine. Measures included students' knowledge of concepts, the numbers and types of analogies teachers used, teacher use of elements of the instructional routine, and teacher and student satisfaction. Teacher use of the routine led to increased student retention and expression of information. In addition, teachers easily learned the routine and used considerably more analogies to instruct concepts after they became familiar with the routine. Teachers indicated that they were satisfied with the routine; students were less satisfied with the routine than were the teachers.

Students and teachers in secondary content classrooms face a variety of challenges. First, the volume and complexity of content information is exploding. Second, today's secondary classrooms are populated with a diverse array of students, including high-achieving and normally achieving students, as well as students who are at risk for failure (e.g., low-achieving students, students with disabilities, students with cultural and economic disadvantages, and students with personal problems). Many of these students lack the skills and strategies needed to meet the demands of secondary content courses, including ones related to understanding and remembering the information covered in those courses (Putnam, Deshler, & Schumaker, 1992). Clearly, instructional methods for teaching abstract and complex information are needed if all students are to benefit from instruction in such courses and meet district and state outcome standards.

Instructional procedures that have the potential of responding to these challenges must be based on theoretically sound principles. If they are to be widely used, they must also be based on sound research studies that incorporate a variety of experimental methodologies and that, taken together, provide convincing evidence that the procedures can yield positive outcomes with regard to student performance in today's classrooms.

One instructional principle that has been discussed in the literature involves the activation of prior knowledge to help students learn new knowledge (e.g., Ausubel, 1963; Brown, Campione, & Day, 1981; Gagne, 1965; Mayer, 1987). Pressley (1995), for example, stressed the importance of encouraging students to relate to-be-learned information to prior knowledge. Berliner (1987) suggested that teachers could help students understand information by making explicit the relationship of new content information to other knowledge possessed by the student. He encouraged teachers either to create the structure for integrating new material with already known material or to prompt students explicitly to be aware of the relevant knowledge they have about the topic being taught.

Vosniadou and Ortony (1989) also held that successful learning often depends on identifying the most relevant bodies of knowledge already in memory so that this knowledge can be used as the starting point for learning something new. To achieve this, they recommended the use of analogies as specific tools to be used to help learners connect new information to prior knowledge. Gentner (1983) indicated that an analogy is "an assertion that a relational structure that normally applies in one domain can be applied in another domain" (p. 156). An analogy has also been defined as a "description or story that tells how two things are similar, even though, on the surface, they do not seem alike" (Bulgren, Schumaker, & Deshler, 1994a, p. 8). Analogies are thought to help learners learn new schema, or complex organizations of information. That is, when given an analogy, students may not need to learn the formal structure of new schema; they simply need to add the new information to already acquired formal structures of knowledge (e.g., Rumelhart & Ortony, 1977). Analogies might be used in content classrooms to (a) foster creativity and stimulate student ability to use knowledge in social studies (Wragg & Allen, 1983), (b) teach a specific chemistry concept (Christian, 1990), (c) teach geography concepts (Andrews, 1987), (d) develop study guides in science (Bean, Singer, & Cowan, 1985), (e) instruct language rules (Connell, 1987), and (f) facilitate literature instruction (McGonigal, 1988).

Nevertheless, the development of analogies for use in content instruction is valuable only if there is reason to believe that students can learn by analogy and that their performance is en-

Janis A. Bulgren, Donald D. Deshler, Jean B. Schumaker, and B. Keith Lenz, Center for Research on Learning, University of Kansas.

Preparation of this article was supported in part by the National Institute of Disability and Rehabilitation Grant H133G90156. We acknowledge the administrators and teachers of Lawrence High School and Central Junior High School in Lawrence (Kansas) School District No. 497, including Ed Judd, Andrea Parson, and Kathy vonEnde, and the administrators and teachers of Shawnee Mission (Kansas) School District No. 512, including Fern Bretch, Peggy Couch, Glenda Johnson, Lynn Martens, Kay Neill, Billye Odom, Harry Roth, Jenny Sands, Dave Taylor, and Kathy Taylor.

Correspondence concerning this article should be addressed to Janis A. Bulgren, University of Kansas Center for Research on Learning, 517 J. R. Pearson Hall, Lawrence, Kansas 66045. Electronic mail may be sent to jbulgren@ukans.edu.

hanced when analogies are presented. Unfortunately, much of the previous research on teaching and learning by analogy has involved the use of laboratory-type studies to describe individuals' ability to reason analogically (e.g., Goldman, Pellegrino, Parseghian, & Sallis, 1982; Sheard & Readence, 1988). Although research has shown that college students can learn to use strategies for solving analogies (Sternberg & Ketron, 1982) and develop a solution to a story problem if they are given an analogous story problem from a different domain (Gick & Holyoak, 1980), that elementary students' problem-solving performance improves following direct instruction in analogical reasoning (White & Alexander, 1984), and that elementary students can transfer understanding of systematic relational structures in a story to similar relationships in other stories (Gentner & Toupin, 1986), no studies have been conducted in relation to adolescent learning. Indeed, no studies have been conducted in secondary content classrooms to determine whether teachers can readily create and use analogies that are tied to their content and whether the use of analogies aids student learning of the complex information contained in secondary content courses. In addition, no data are available on the effects of analogy use on the performance of the very students whose learning needs to be enhanced—those who are at-risk for failure in secondary content classes.

Thus, the purpose of this project was to explore the use of analogies as an instructional tool to teach important, difficult concepts (e.g., imperialism, commensalism, and personification) in secondary content classrooms containing diverse groups of students. A combination of quantitative, qualitative, single-subject, and large-group methodologies was selected to provide information about analogies as an instructional technique. Study 1 was conducted to determine the effects of an analogically based instructional routine on students' knowledge of concepts. This study provided evidence that the instructional routine had the power to affect student learning. Study 2 was conducted to determine the effects of instruction in the routine on 10 secondary science and social studies teachers' use of the analogically based routine in their classes. This study was conducted as a logical step in determining whether the procedures validated in a controlled setting could be integrated into regular content instruction by a variety of teachers and whether teachers and students would be satisfied with the routine. Study 3, which was conducted in an intact secondary classroom setting, was conducted to determine student knowledge of important science concepts that the teacher taught with and without the use of the analogically based routine.

These studies were part of a line of programmatic research focusing on instruction in secondary-level subject-matter classes that contain students of diverse abilities (Bulgren & Lenz, 1996; Lenz, Bulgren, & Hudson, 1990; Schumaker, Deshler, & McKnight, 1991). Overall, the main goal associated with this line of research has been to utilize a wide variety of research methods to determine the effectiveness of a series of instructional routines designed to enhance student performance. The current group of studies on utilizing analogies to teach important and difficult information illustrates this research philosophy in that it includes initial cooperation between teachers, researchers, and administrators to identify the needs of a district; development of an instructional routine that fits those needs through ongoing consultation with a cadre of expert teachers and outside consultants (Bulgren, Lenz, Scanlon, & Clark, 1995); validation of the power of the instructional routine to enhance learning by students of varied abilities in controlled experiments; validation of teacher-training procedures and observation of teachers by researchers in secondary content classrooms to assure the applicability of the instructional routine in a variety of subject-area courses with real classroom curricula; collection of student data resulting from implementation of the instructional techniques by teachers in intact classrooms to validate learning in actual classroom settings under real-world conditions; collection of satisfaction data from both teachers and students following the use of the instructional routine; and ongoing cooperative refinement of the routine with teachers and researchers after the initial studies.

The focus of the three studies was an instructional device called the Concept Anchoring Table and an instructional routine called the Concept Anchoring Routine (Bulgren et al., 1994a), which was coconstructed with school personnel in the participating school district (see Lenz, Schumaker, & Deshler, 1991, for a description of this process). The device and routine are based on principles associated with content enhancement (Bulgren & Lenz, 1996; Lenz et al., 1990; Schumaker et al., 1991), which is a process of teaching scientific or cultural knowledge to heterogeneous classes of students in which both group and individual learning needs are met; the integrity of the content is maintained; critical content is selected, organized, manipulated, and complemented in a manner that promotes effective and efficient information processing; and the content is delivered in partnership with students in a manner that facilitates and enriches learning for all students.

The Concept Anchoring Table (Figure 1) was designed to provide teachers with a teaching tool for visually displaying an analogy that connects known information (i.e., prior knowledge) with new information and that promotes uniformity of instructional use of analogies across teachers and across teaching instances. The table is a one-page graphic device with places for writing information related to two concepts: a "known concept" and a "new concept." Through the use of the table, teachers can enhance student understanding of a new, difficult concept by relating key characteristics of that concept to similar characteristics of a known or familiar concept. For example, in Figure 1, the known concept is "temperature control systems in modern buildings," and the new concept is "temperature control systems in warm-blooded animals." Characteristics of each of the concepts are listed in the sections under the concepts so that parallel characteristics are next to each other; in the center section, space is provided to write the shared characteristics, that is, statements of how the parallel characteristics shared by the new concept and the known concept are similar. The characteristics of the known concept and their organization provide an analogy for the characteristics of the new concept to be learned. At the bottom of the table is a box where a statement of understanding about the new concept is written.

An initial draft of a Concept Anchoring Table is to be completed by a teacher for each new concept to be taught prior to presenting the analogy and the new concept in class. Students are to receive a blank form of the table, and the final version of the table is coconstructed by the teacher and students in partnership through the use of the Concept Anchoring Routine.

The Concept Anchoring Routine was developed on the basis of previous research related to concept instruction (Bulgren, Schumaker, & Deshler, 1988) as well as teacher input and includes a

Anchoring Table

Unit: ________ Name: ________ Date: ________

③ **Known Information**

furnace
air conditioner
72 degrees
thermostat
notices change
set temp. anywhere
electronic signals

② Known Concept		① New Concept
Temperature control systems in modern buildings		Temperature control systems in warm-blooded animals
④ Characteristics of Known Concept	**⑥ Characteristics Shared**	**⑤ Characteristics of New Concept**
Building temperature is set to stay the same (72 °F). →	Inside temperature is supposed to stay the same.	← Body temperature stays the same (98.6 °F).
Thermostats notice temperature changes. →	Something notices temperature changes.	← Nervous & endocrine systems notice temperature changes.
When temperature changes, thermostat sends electronic signals. →	When temperature changes, a sensor sends signals.	← When temperature changes, nervous & endocrine systems send signals.
Signals start action in furnace or air conditioner. →	Signals start other systems.	← Signals start action in circulatory system or muscles.
Furnace or air conditioner corrects building temperature to 72 °F. →	Systems correct temperature.	← Circulatory system & muscles correct body temperature.

⑦ **Understanding of the New Concept:** Temperature control systems in warm-blooded animals are like those in modern buildings because the temperature is supposed to stay the same, but when the temperature changes, something notices. A sensor sends signals to start other systems that correct the temperature.

ANCHORS Linking Steps:	1 Announce the New Concept	2 Name Known Concept	3 Collect Known Information	4 Highlight Characteristics of Known Concept	5 Observe Characteristics of New Concept	6 Reveal Characteristics Shared	7 State Understanding of New Concept

Figure 1. Example Concept Anchoring Table for the concept "temperature control systems in warm-blooded animals." Anchoring Tables are data-based teaching instruments that have been found to be effective when used with a teaching routine that combines cues about the instruction, specialized delivery of the content, involvement of the students in the cognitive processes, and a review of the learning process and content material as described in this article. They have not been shown to be effective tools if simply distributed to students. From *The Concept Anchoring Routine* (p. 34), by J. A. Bulgren, J. B. Schumaker, and D. D. Deshler, 1994, Lawrence, KS: Edge Enterprises, Inc. Copyright 1994 by Janis A. Bulgren, Jean B. Schumaker, and Donald D. Deshler. Reprinted with permission.

number of validated teaching methods such as advance organizers (e.g., Ausubel, 1963; Lenz, Alley, & Schumaker, 1987) and interactive devices (e.g., Markman, 1985; Palincsar & Brown, 1984; Raphael & Gavalek, 1984; Wong, 1985). The name of the routine is based on the notion that the use of an analogy anchors new knowledge to prior knowledge.

The Concept Anchoring Routine has three instructional phrases: cue, do, and review. During cue, students are informed that the routine will be used and are given instructions on how to participate so as to assure active student learning. Specifically, the cue part of the routine involves the teacher (a) indicating the importance of the new concept, (b) explaining the benefits of understanding the concept, (c) explaining that a graphic device will be used to promote understanding of the concept, and (d) prompting students to participate in the coconstruction of the graphic device and to take notes about the concept.

The do component of the Concept Anchoring Routine involves the coconstruction of the Concept Anchoring Table. The instructional sequence follows the numbers in the sections on the Concept Anchoring Table (see Figure 1) and includes naming the new concept, naming the known concept, identifying information about the known concept and listing important characteristics of the known concept, listing parallel characteristics of the new concept, naming the characteristics shared by both the new and the known concepts, and creating a summary statement such as a definition of the new concept or a statement about why the two concepts are comparable. The teacher completes the first two steps, whereas the remaining steps are completed by the students and teacher in partnership in a discussion format.

To promote the discussion, the teacher uses probing questions based on the initial draft of the Concept Anchoring Table, which was made prior to class; however, because the coconstructive process involves the collaborative construction of meaning, the final version of the table is often somewhat different from the initial draft. The teacher also uses organizing statements (e.g., "Let's begin by . . ."), statements of rationale (e.g., "The reason why we are brainstorming ideas first is . . ."), cues regarding participation and expectations (e.g., "Please take notes on the Concept Anchoring Table and keep it in your portfolio for reference as we study vertebrates."), questions to check student understanding (e.g., "What is one characteristic that makes temperature regulation in warm-blooded animals similar to the thermostat in your home?"), and feedback statements (e.g., "That's a good point.") throughout the discussion.

In the review component of the Concept Anchoring Routine, the teacher and students review all content written in the table and the process used in creating the analogy and the Concept Anchoring Table. To do this, the teacher asks the students questions (e.g., "What is one important characteristic of a temperature-control system?") and process questions (e.g., "How does making connections between temperature-control systems in modern homes and temperature-control systems in warm-blooded animals help you as a learner?"). The purpose of the review component is to check student understanding and retention of the information as well as understanding of the cognitive processes used to generate the table. The review can be conducted any time after the initial presentation of the table, including during the same class period, the next day, and the day of the review for an upcoming test.

Study 1: Effects of the Routine in Specially Designed Lessons

Method

Participants and Settings

Eighty-three students were recruited from the general education classes of three teachers in three high schools in a midwestern United States suburban school district. All participating students were enrolled in a course titled Introduction to Investigative Science Skills. The study took place in three typical general education classrooms, each of which had desks, chairs, an overhead projector, and a screen. Two teachers taught the course in two classes each; the third teacher taught the course in four science classes. The eight classes were randomly assigned to one of two experimental conditions, hereafter referred to as Condition 1 and Condition 2. The study took place during regularly scheduled classes. For the selection process, students volunteered to allow their data to be used in this study by returning consent forms signed by their parents. Although all students present in the classrooms received the instruction, 39 students who participated in Condition 1 and 44 students who participated in Condition 2 had permission for their data to be used.

Four types of students were included: students classified for the study as high achievers (HA), normal achievers (NA), low achievers (LA), and students with learning disabilities (LD). The HA students had grade point averages (GPAs) of 3.50 and above. The NA students had received no more than one grade below the "C" level in either semester of the previous school year and had GPAs of 3.49 or lower. The LA students had received at least two grades below the "C" level in academic courses during at least one of the two semesters of the previous school year. Students with LD had been formally classified as such by their school districts, which followed district and state guidelines. Twelve HA students (6 in Condition 1 and 6 in Condition 2), 28 NA students (14 in Condition 1 and 14 in Condition 2), and 15 LA students (8 in Condition 1 and 7 in Condition 2) were involved in the study.[1]

Information was also collected on students with LD. For the 11 students with LD in Condition 1, their mean full scale standard score on the Wechsler Intelligence Scale for Children—Revised (WISC-R) was 99.00 (SD = 13.35, range = 74–116); their mean national percentile score on the Woodcock-Johnson Written Language Test was 30.18 (SD = 23.68, range = 8–74); and their mean national percentile score on the Woodcock-Johnson Reading Test was 29.09 (SD = 16.99, range = 11–66). For the 17 students with LD in Condition 2, their mean standard full scale score on the WISC-R was 100.82 (SD = 11.35, range = 78–116), their mean national percentile score on the Woodcock-Johnson Written Language Test was 33.53 (SD = 19.99, range = 12–92), and their mean national percentile score on the Woodcock-Johnson Reading Test was 37.00 (SD = 21.09, range = 2–63).

The Content Lesson

The topic for the content lesson, "relationships in the environment," was selected because (a) a lesson about the topic could be designed that contained the type of information that students in an inclusive science content classroom might be expected to understand and remember, (b) the three teachers agreed that content associated with the topic was relevant to their courses, (c) the teachers concurred that the topic had not been covered in their courses, and (d) the teachers predicted that their students would have limited prior knowledge about the topic.

Four concepts associated with the topic were targeted for instruction. Two concepts were chosen for analogical instruction: *pyramid of numbers* and *commensalism*. For pyramid of numbers, the concept of *military structure* was selected as the known concept, and the analogy was built on this known concept because the teachers indicated that students were familiar with it due to current events involving military operations. In addition, the new concept and the known concept share a structure in which the numbers of living things are large at lower levels of the relationship structure and small at the top levels. To teach the concept of commensalism, an analogy consisting of a story about human relationships in a lemonade-stand business was constructed because the teachers indicated that they would use stories or situations as they taught students about concepts such as commensalism. The new concept, commensalism, and the known concept, a lemonade-stand business, both illustrate a relationship between two living things in which one of the living things benefits, but in which the other living thing is neither benefited nor harmed.[2] Two other concepts were selected for traditional instruction: *food web* and *heterotroph*.

Two 35-min scripts were constructed for the content lesson about the four concepts: one for Condition 1 and one for Condition 2. The two scripts

[1] Complete demographic information on the students is available on request from Janis A. Bulgren.

[2] The story about the lemonade stand is as follows: A boy wants to put his lemonade stand on the corner of his neighborhood street because he thinks he will get more business there than in the middle of the block where he lives. He asks the woman who lives on the corner lot if he may set up his lemonade stand on the sidewalk on the corner. He promises to clean up his trash before he goes home. She gives permission to set up the stand, and he follows through on his promise. Thus, the boy and the woman have a relationship/agreement. The boy is benefited, and the woman is neither benefited nor harmed. This relationship of the boy and the woman is similar to the relationship between Spanish moss and the tree that the moss lives in. The moss benefits from living high in the tree branches because it can receive sunlight. The tree is neither benefited nor harmed. Both relationships are examples of commensalism. The three common characteristics that would be highlighted here are that (a) both relationships involve two parties, (b) one party is benefited, and (c) one party is neither benefited nor harmed.

were identical in certain respects and different in others. Both scripts contained the same presentation of logically ordered information about relationships in the environment. Both included the same information about the four concepts. In addition, at the beginning of both scripts were statements naming the topic, giving rationales regarding the benefits of understanding the topic, describing the importance of understanding the information to be presented about the topic, prompting students to take notes, and explaining that the instruction would help them learn. Other similarities between the two scripts were that they both contained a prompted discussion about the environment of the classroom to establish prior knowledge about environmental relationships and a reminder that students would take a test on the material on the next day. In addition, both scripts contained the same number of questions to be used to elicit student participation.

The two scripts were different only during the portion of the presentation in which one of the four concepts was to be presented with the use of the Concept Anchoring Table and the Concept Anchoring Routine. In the script for Condition 1, the concept of commensalism was associated with the table and routine; in the script for Condition 2, the concept of pyramid of numbers was associated with the table and routine. (Hereafter, the concept associated with the routine is referred to as the "enhanced" concept.)

Words to be written on the overhead projector appeared in boldface print throughout both scripts to prompt the instructor to write items associated with critical relationships on the overhead projector. Information to be listed included the topic, "relationships in the environment," and the following clusters of information defining relationships in the environment: (a) environment—a setting composed of living and nonliving things; (b) mutualism—a relationship in which both living things benefit; (c) competition—a relationship in which both living things are harmed; (d) amensalism—a relationship in which one living thing is neutral and another is harmed; (e) commensalism—a relationship in which one living thing is helped and the other is neither helped nor harmed; (f) autotrophs—green plants or producers of their own food; (g) heterotrophs—animals, or consumers of food, who cannot make their own food; (h) food chain—a sequence of living things containing producers, consumers, and decomposers; (i) the pyramid of numbers—a way to explain the decrease in energy and numbers of living things at each successively higher level in the food chain; (j) decomposers—living things that break down once-living, decaying matter; and (k) food web—more than one food chain. In addition, all of the characteristics related to the concepts of commensalism, pyramid of numbers, food web, and heterotroph were also boldfaced in the script and were to be written on an overhead transparency during both lessons. The only difference regarding what was to be written during the lessons related to the characteristics of the known concept to be associated with the enhanced concept. During the Condition 1 lesson, the instructor was to write the characteristics related to the lemonade-stand business and the shared characteristics between commensalism and the lemonade-stand; during the Condition 2 lesson, the instructor was to write the characteristics related to military structure and the shared characteristics between pyramid of numbers and military structure.

Measurement System

A 32-item multiple-choice test was used to measure recognition of facts and understanding of four concepts that were included in the lesson (commensalism, pyramid of numbers, food web, and heterotroph). Eight questions pertaining to each concept were included. The following is a sample test item: What term defines a relationship in which Organism A is helped by Organism B, but Organism B is neither helped nor harmed? (1 = commensalism, 2 = mutualism, 3 = amensalism, 4 = competition).

Item difficulty index scores were computed on each item by administering an 80-item test to 8 junior high school students attending a different school than the study participants. These students received a nonenhanced lesson on the four concepts (i.e., the information on all four concepts was taught without the Concept Anchoring Table and routine). After the students had taken the test and their responses had been scored as correct or incorrect, an item-difficulty index was computed for each item by determining the proportion of students responding correctly to it. The index was used to select the 32 items of equivalent difficulty to be included in the test. Internal consistency for the resulting instrument was calculated using Cronbach's alpha to estimate internal consistency. An alpha score of .92 was obtained for the total test.

In addition, the test items were analyzed by four judges. Two of the judges were certified teachers in the area of science. One held a bachelor's degree, had over 60 hr of graduate study, and had extensive experience in teaching and curriculum; the other held a master's degree in biology and had 36 hr of graduate study in secondary science education. The third judge held a doctorate degree in special education and was certified to teach at the secondary level. The fourth judge had expertise in test construction and assessment and was pursuing a graduate degree in educational psychology and research. Content validity for the instrument was established by having each judge (a) associate each item with one of the four concepts and (b) match each item on the test to information in the script. All members of the panel determined that the same eight items were associated with each concept and that every item on the test was covered in the script.

Interscorer reliability on the student test was determined by having two scorers independently score a random sample (15%) of the tests taken by students assigned to both conditions. The two scorers recordings were compared item by item, and 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 763 times out of 768 opportunities to agree, for a percentage of agreement of 99% (the range on individual tests was 96% to 100%).

Procedures

The instructor was Janis A. Bulgren, who holds a doctorate degree in special education and has been certified to teach English at the secondary level. During each class, she made every effort to deliver the material naturally, while following the appropriate script for the designated condition. A second researcher was present for all classes to assure procedural integrity. The sessions were audiotaped, and the second researcher recorded information delivered by the instructor on a checklist of items corresponding to the information in the script and took notes, replicating all the information that the instructor wrote on the overhead transparency. A third researcher listened to the tapes and used the same checklist to record the information presented by the instructor. The third researcher's recordings agreed with the second researcher's recorded observations 100% of the time. Tapes and written notes were also reviewed to verify that all of the critical information was presented in all conditions; it was. Each class period lasted 45 min; each lesson lasted 35 min. Ten minutes were spent in classroom management activities. Within each lesson, the same amount of time was spent on instruction in both conditions for instructions to, and general information about, the topic of relationships in the environment and on each targeted concept. Notebook paper was distributed to all students at the beginning of each lesson, and the students were instructed to take notes throughout the lesson. Student notes were collected at the end of the lesson.

Instruction was delivered as specified in the script for each lesson. Instruction on the enhanced concepts included the use of the Concept Anchoring Routine and coconstruction of the Concept Anchoring Table. The traditional lecture-style instruction included oral teacher presentation of the information in the script, written presentation of information on an overhead transparency at times specified in the script, and interactive discussion with the students utilizing the questions specified in the script.

On the next day, the test was administered. Test instructions were read aloud to the students; test items were not read to the students. Students in both conditions were given the same test. Students were given 35 min to complete the test. They were not allowed to consult their notes or each other as they took the test.

Experimental Design

The experimental design is shown in Table 1. The students' performance on items related to the enhanced concept in each condition was compared with the students' performance on items related to the same concept when it was presented through the use of the traditional lecture–discussion format in the other condition to determine the effects of the Concept Anchoring Routine. Two different concepts were enhanced, one in each condition; therefore, each group of students acted as a control for the other. In addition, the students' performances on items related to the two concepts that were never associated with the routine (food web and heterotroph) were used as further controls in the study.

Results

Student Tests

Whole-group results are shown in Figure 2.[3] A univariate analysis of variance indicated that Condition 1 students (M = 77%, SD = 25.17), who received the enhanced explanation of commensalism, answered significantly more commensalism items, $F(1, 81) = 20.03$, $p = .000$, than did Condition 2 students (M = 52%, SD = 24.85), who received traditional instruction on that concept. In addition, Condition 2 students, who received the enhanced explanation of pyramid of numbers, answered significantly more pyramid of numbers items (M = 80%, SD = 19.14) than did Condition 1 students, who received traditional instruction on that concept (M = 64%, SD = 30.32), $F(1, 81) = 9.12$, $p = .001$). There were no significant differences between the groups with regard to their scores on items related to heterotroph, $F(1, 81) = .294$, $p = .589$, and food web, $F(1, 81) = .772$, $p = .383$. A Bonferroni correction indicated no differences in the results.

The mean test scores of students in the subgroups (HA, NA, LA, and LD) are shown in Figure 3. Because of the small sample size in the subgroups, the Kruskal-Wallis test was used to determine whether the difference in performance by each subgroup of students was significant across the conditions. When commensalism was enhanced versus not enhanced, the difference was significant at the .05 level for LAs (enhanced: M = 80%, SD = 24.03; nonenhanced: M = 46%, SD = 27.68; $p = .03$), NAs (enhanced: M = 84%, SD = 19.26; nonenhanced: M = 64%, SD = 19.52; $p = .007$), and HAs (enhanced: M = 96%, SD = 6.46; nonenhanced: M = 75%, SD = 13.69; $p = .007$). The difference between the performances of the students with LD in the two conditions (enhanced: M = 55%, SD = 25.78; nonenhanced: M = 36%, SD = 19.71) did not reach significance ($p = .051$).

When pyramid of numbers was enhanced versus not enhanced, the difference was significant at the .05 level for LDs (enhanced: M = 69%, SD = 15.38; nonenhanced: M = 40%, SD = 24.89; $p = .002$) and NAs (enhanced: M = 92%, SD = 19.52; nonenhanced: M = 73%, SD = 24.93; $p = .02$). The difference for the LA students did not reach significance (enhanced: M = 73%, SD = 20.95; nonenhanced: M = 53%, SD = 24.78; $p = .11$), nor did the difference for the HA students (enhanced: (M = 94%, SD = 10.46; nonenhanced: M = 100%, SD = 0.00; $p = .14$). The Kruskal-Wallis test indicated that there were no differences in average student performance across the conditions for any of the subgroups for heterotroph (LDs: $p = .944$; LAs: $p = .685$; NAs: $p = .872$; HAs: $p = .522$;) or for food web (LDs: $p = .240$; LAs: $p = .643$; NAs: $p = .395$; HAs: $p = .471$).

A further analysis was conducted to determine the percentage of students who performed at a level commonly deemed as "passing" in secondary content classes (i.e., who earned a score of 60% or above) on parts of the test. Students' scores on the commensalism test items represent passing grades (i.e., scores above 60%) for the following percentage of students: 36% of the LDs in the enhanced condition compared with 12% in the nonenhanced condition; 75% of the LAs in the enhanced condition compared with 29% in the nonenhanced condition; 93% of the NAs in the enhanced condition compared with 64% in the nonenhanced condition; and 100% of the HAs in the enhanced condition compared with 83% in the nonenhanced condition.

Students' scores on the pyramid of numbers test items represent passing grades for the following percentages of students: 77% of the LDs in the enhanced condition compared with 27% in the nonenhanced condition; 86% of the LAs in the enhanced condition compared with 50% in the nonenhanced condition; 93% of the NAs in the enhanced condition compared with 71% in the nonenhanced condition; and 100% of the HAs in the enhanced condition compared with 100% in the nonenhanced condition.

Data were also analyzed to determine the percentage of students functioning at the 75% level, a level often associated with functioning at the middle "C" level. When commensalism was enhanced, the same percentages of students passing at the 60% level also earned scores at or above the 75% level in the LD, LA, NA, and HA groups; that is, 36% of the LD, 75% of the LA, 93% of the NA, and 100% of the HA students earned scores at or above this level. However, when commensalism was not enhanced, 12% of the students with LD, 29% of the LA students, 36% of the NA students, and 83% of the HA students earned scores at or above the 75% level.

When pyramid of numbers was enhanced, 47% of the students with LD, 57% of the LA students, 87% of the NA students, and 100% of the HA students earned scores at or above the 75% level. When pyramid of numbers was not enhanced, 9% of the students with LD, 25% of the LA students, 64% of the NA students, and 100% of the HA students earned scores at or above the 75% level.

Table 1
Design for Study 1

	Subgroups of students							
	Condition 1				Condition 2			
Concept	LD	LA	NA	HA	LD	LA	NA	HA
Pyramid of numbers	E	E	E	E	NE	NE	NE	NE
Commensalism	NE	NE	NE	NE	E	E	E	E

Note. LD = students with learning disabilities; LA = low achievers; NA = normal achievers; HA = high achievers; E = enhanced; NE = not enhanced.

[3] The data from Study 1 were analyzed using individual student scores rather than class means because differences in the way classes were constituted made individual scores more meaningful than class means.

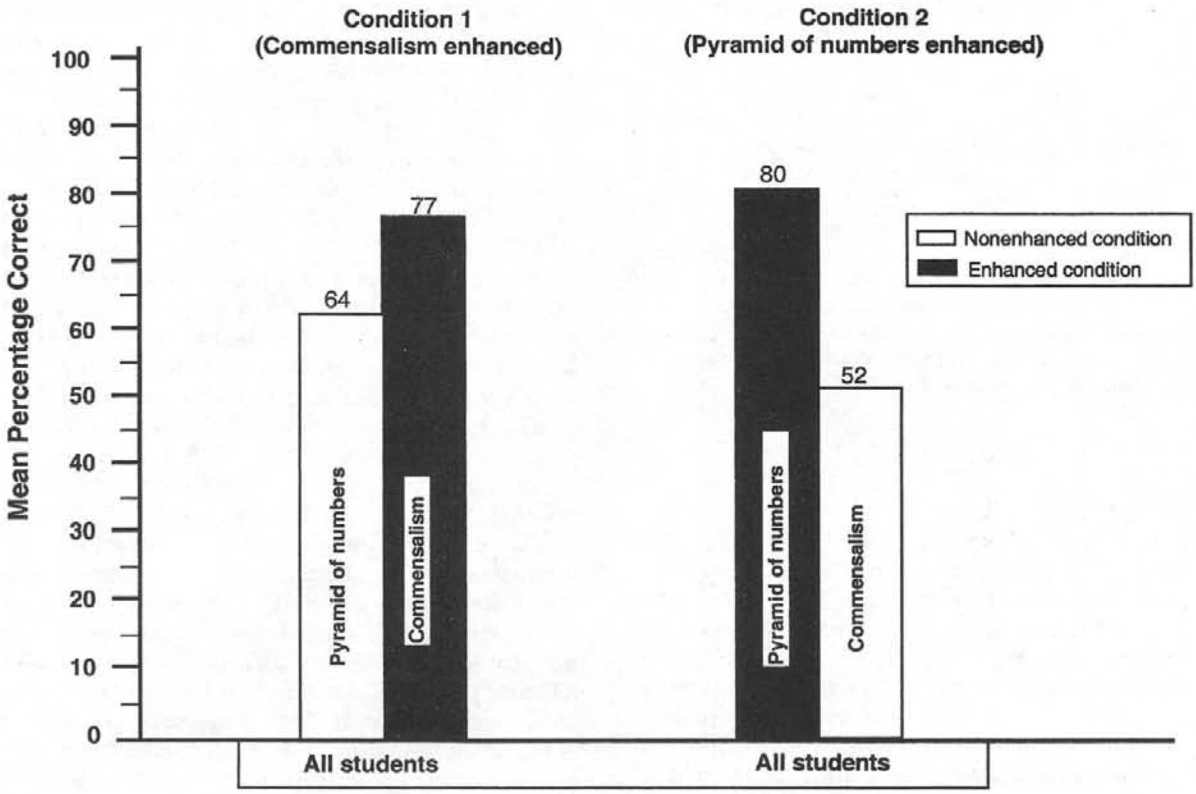

Figure 2. Mean percentage scores earned by all students in Conditions 1 and 2 on a test designed to assess knowledge of concepts in enhanced and nonenhanced conditions.

Conclusions

In summary, the use of the Concept Anchoring Routine to teach difficult concepts results in significantly better student performance on tests. Furthermore, although the differences were not significant for all the subgroups in every case, the routine sometimes helped high-achieving and normally achieving students, as well as at-risk students (LAs and LDs), understand and remember information better. Because the routine appeared promising, Study 2 was conducted to determine whether teachers could and would integrate the routine within their ongoing instruction in secondary content courses and to determine teachers' and students' reactions to the routine.

Study 2: Effects of Training on Teacher Use of the Routine

Method

Participants

Teachers. Ten secondary content teachers who taught in two school districts located in suburban areas of eastern Kansas volunteered to participate in Study 2. They received $80 each for their participation, which lasted approximately 8 months. Descriptive information about the teachers, their classes, and their schools is shown in Table 2.

Students. Each teacher targeted one class of students for participation. There were 193 students in the 10 targeted classes who supplied satisfaction data about the instructional methods used by the teachers.

Settings

The classrooms of the participating teachers served as the settings. They were typical classrooms with desks, chairs, chalkboards, and overhead projectors. The study took place during regularly scheduled class periods that ranged from 45 to 55 min in length.

Measurement Systems

Implementation Checklist. A 12-point checklist (hereafter referred to as the Implementation Checklist) was used to assess teacher implementation of the Concept Anchoring Routine in their classrooms. It listed items corresponding to each of the steps in the Concept Anchoring Routine, which were outlined earlier. Next to each item was a space where an observer could record points earned by the teacher for completing the corresponding step of the routine.

Each item on the checklist was objectively defined in writing in an evaluation manual, written examples were provided, and the item was assigned a point value ranging from 3 to 15 points. For example, a teacher received 3 points for cueing the students to take notes about a concept, 10 points for completing each step in the "Do" part of the routine, and 15 points for using a visual depiction of the information. If a step was omitted, the teacher received 0 points for that step. No partial credit was awarded.

Observers were instructed to begin recording a teacher's behavior on the checklist any time a teacher indicated a concept was important, explained the benefits of understanding a concept, or presented a way to help students understand a concept (hereafter, these instances are referred to as "opportunities to use the routine"). A total of 100 points could be earned by each teacher each time an opportunity to use the routine occurred. A percentage score was calculated for each teacher's performance. If the teacher cued the

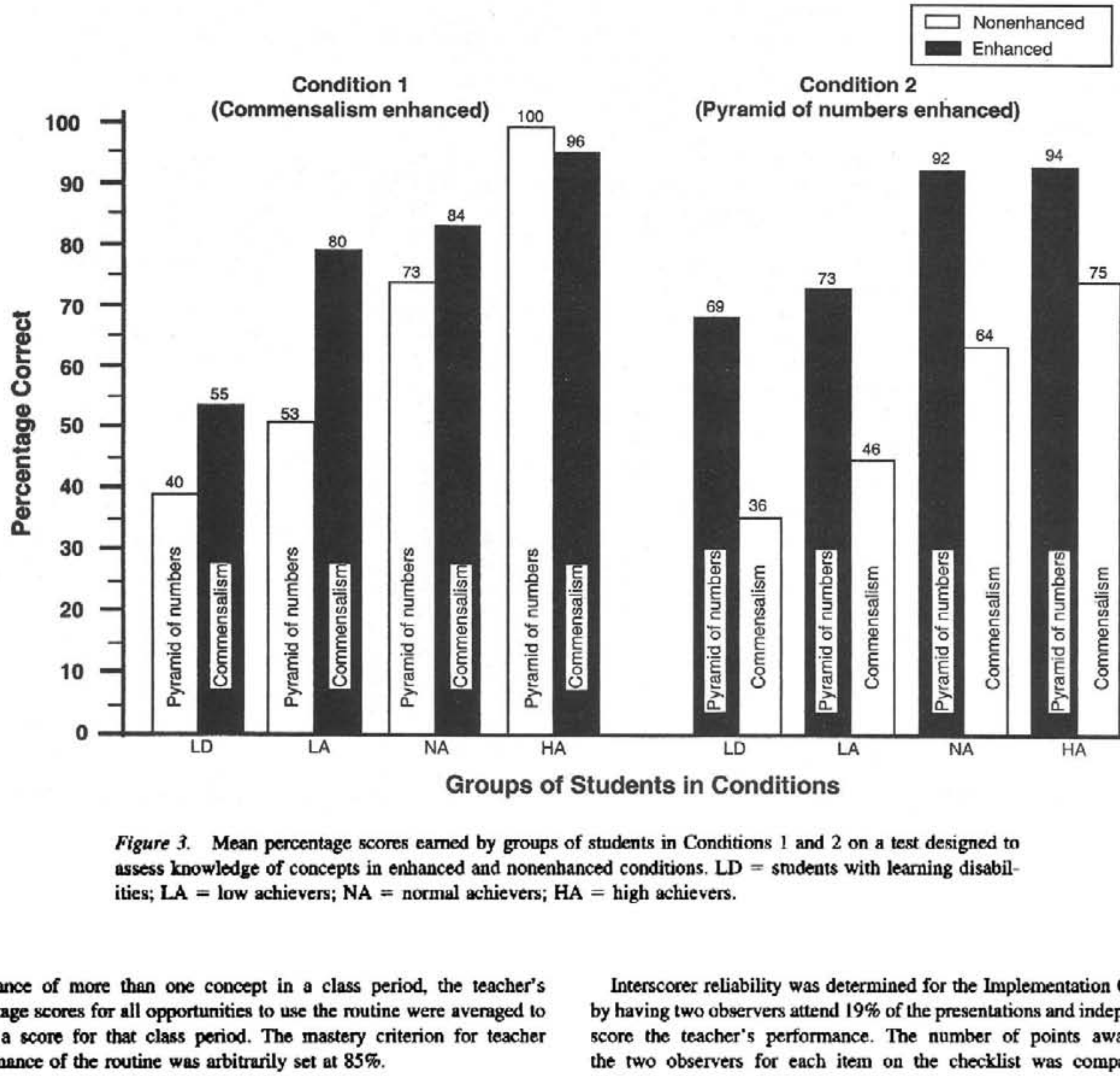

Figure 3. Mean percentage scores earned by groups of students in Conditions 1 and 2 on a test designed to assess knowledge of concepts in enhanced and nonenhanced conditions. LD = students with learning disabilities; LA = low achievers; NA = normal achievers; HA = high achievers.

importance of more than one concept in a class period, the teacher's percentage scores for all opportunities to use the routine were averaged to obtain a score for that class period. The mastery criterion for teacher performance of the routine was arbitrarily set at 85%.

Interscorer reliability was determined for the Implementation Checklist by having two observers attend 19% of the presentations and independently score the teacher's performance. The number of points awarded by the two observers for each item on the checklist was compared. An

Table 2
Descriptive Information About the Teachers and Schools

Teachers	School district	No. students served in school	Grades taught	Subject	Age	Gender	Race	Years of experience
1	A	330	7	Life science	48	M	C	23
2	B	1,750	10	Biology	49	M	C	29
3	A	700	7	Geography	46	F	C	25
4	B	550	7	Social studies	43	F	C	20
5	A	330	7	American history	44	F	C	23
6	A	1,700	10–12	Chemistry	41	F	C	14
7	A	700	7	Life science	35	F	C	11
8	A	950	7	Life science	48	F	C	17
9	A	1,700	9	Geography	51	F	C	16
10	B	550	7	Geography	50	F	C	11

Note. Teachers 1–10 participated in Study 2; Teacher 8 participated in Studies 2 and 3. The means for teachers' age and years of experience were 46 and 19, respectively. A = District A; B = District B; M = male; F = female; C = Caucasian.

agreement was scored if both observers recorded the same number of points for 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. There were 149 agreements within 156 opportunities for agreement, for a percentage of agreement of 96% (range = 67% to 100%).

Types of analogies used. Observers also used a score sheet on which types of analogies were defined and several written examples of each type were provided. An analogy was defined as any description or story that tells how two or more things are similar.

Each analogy was categorized according to three sets of descriptors on the score sheet. Scorers simply placed a checkmark next to the most appropriate descriptors in each set. First, an analogy was identified as either a between-domains analogy (sometimes called a "metaphorical analogy") or a within-domain analogy (sometimes called a "literal analogy") (Vosniadou & Ortony, 1989). A between-domains analogy was defined as a comparison in which the known concept (e.g., sponge) and the new concept (e.g., alveoli) were not in the same domain (i.e., cleaning implements and body parts), whereas a within-domain analogy was a comparison in which both the known concept (e.g., school counselors) and the new concept (e.g., genetic counselors) were in the same domain (i.e., the profession of counseling).

Second, each known concept used in an analogy was identified as either abstract or concrete (Holyoak, 1984). Abstract concepts were those such as family relationships and independent living, whereas concrete concepts were those such as layered jello and a swinging pet door.

Third, each analogy was identified as showing a structural relationship between two concepts, a functional relationship, or both (Curtis, 1988). A structural relationship was scored if the analogy was used to point out how the two concepts were structurally similar (e.g., a ladder and DNA); a functional relationship was identified if the analogy was used to point out how the two concepts were functionally similar (e.g., respiration and fermentation).

Interscorer reliability on the types of analogies created by the teachers was determined by having two scorers independently score all of the analogies used by the teachers after training. An agreement was scored if both scorers indicated the same descriptor for a given analogy. The percentage of agreement was calculated as described earlier; it was 100% (99 agreements within 99 opportunities for agreement).

Teacher Satisfaction Questionnaire. The Teacher Satisfaction Questionnaire contained 20 items constructed as 7-point Likert-type scales ranging from 1 (*completely satisfied*) to 7 (*completely dissatisfied*) or from 1 (*very likely*) to 7 (*very unlikely*). Items related to such topics as the flexibility of the routine, acceptability of preparation time, ease of use of the routine, student learning, usefulness of the Concept Anchoring Table, student attention, student note taking, and student use of the table to study for tests. In addition, the questionnaire explored whether teachers would continue to use the routine in their classes, would recommend it to other teachers, and would recommend that other teachers learn about the routine if inservice instruction were available.

Student Satisfaction Questionnaire. Satisfaction ratings were also elicited from students. Items on the Student Satisfaction Questionnaire related to how satisfied the students were that the Concept Anchoring Table helped them follow what the teacher was saying, take notes, focus their attention on the information, study for tests, do well in class, and improve their grades in the class. They were also asked to indicate how satisfied they were with this new way of teaching compared with when the teacher did not use the Concept Anchoring Table.

Procedures

Teacher instruction. Two researchers presented information about the Concept Anchoring Table and the Concept Anchoring Routine to the teachers in a 2-hr workshop session that included a description of the table and the routine, rationales for using analogies within the instructional process, a demonstration of the routine, practice developing Concept Anchoring Tables, practice using the routine, individual feedback on development of the tables and performance of the routine in simulated practice sessions, and group discussion.

In addition, the teachers were given a written description of the Concept Anchoring Table and the routine in the form of a guidebook (Bulgren et al., 1994a). They were encouraged to use the guidebook as a reference during the study. The guidebook included rationales for using analogies within instruction, definitions associated with concept anchoring, descriptions of types of information that can be taught with concept anchoring, a description of the Concept Anchoring Table, guidelines for preparing Concept Anchoring Tables, examples of Concept Anchoring Tables, an explanation about how to introduce the Concept Anchoring Table to students, a description of how to coconstruct Concept Anchoring Tables with students, and guidelines for evaluating how well students have learned the targeted concepts.

Implementation of the routine. After the 2-hr inservice session, teachers implemented the routine in their classes as often as they wished. They were asked to choose at least two concepts that they wanted students to learn, to prepare a draft of the Concept Anchoring Table for each concept, and to use the Concept Anchoring Routine to coconstruct a final version of the Concept Anchoring Table with students in the classroom. Researchers consulted with teachers individually, as requested by the teachers, as they planned their Concept Anchoring Tables and their presentations of the Concept Anchoring Routine.

Experimental Design

A multiple-baseline across-teachers design (Baer, Wolf, & Risley, 1968) was used with 2 teachers to determine the effects of the workshop instruction on teacher behavior in the classroom. Then the design was replicated four times, with 2 teachers taking part in each replication. During baseline, the teachers were observed during at least three class sessions in which they indicated that they would be presenting concepts that they expected the students to understand and remember. Their behavior was recorded on the Implementation Checklist.

After a stable baseline was achieved for the 1st teacher in each multiple-baseline design, the teacher received instruction in how to use the routine, as described previously, in an inservice session and began implementing the routine. After this teacher substantially improved over his or her baseline performance, as indicated by the score on the Implementation Checklist, the 2nd teacher in the design received instruction on how to use the routine and began implementing it. Observers continued visiting the classrooms and recording teacher behavior throughout the remainder of the school year, whenever the teachers indicated they would be using the routine. Satisfaction questionnaires were administered to the teachers and the students at the end of the year.

Results

Teacher Performance of the Routine

The performances with regard to implementing the Concept Anchoring Routine are summarized in Figure 4 for Teachers 1–6 and in Figure 5 for Teachers 7–10. The percentage of points earned on the Implementation Checklist by teachers is shown. For each teacher, baseline performances are shown to the left of the vertical line in each graph, and post-training performances are shown to the right of the line.

During baseline, teacher scores on the Implementation Checklist ranged from 0% to 40% (M = 4%). In general, before training, teachers occasionally received points for cueing the students that a concept was important, for explaining the benefits of understand-

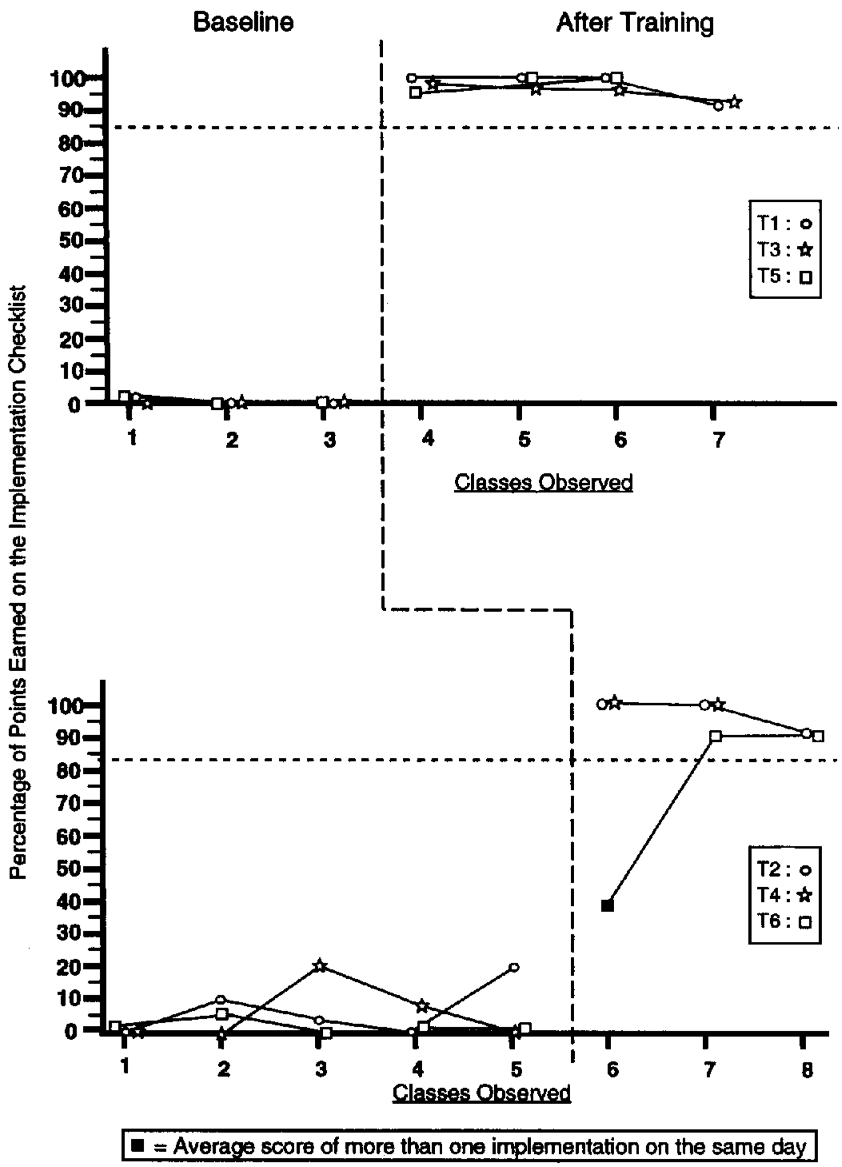

Figure 4. Percentage of points earned on the Implementation Checklist by Teachers (T) 1–6.

ing the concept, or for presenting characteristics of new and familiar concepts. For example, a total of 14 cues about the importance of a concept were observed in 11 of the 38 baseline observations. Across all 38 baseline observations, the teachers provided the students with an analogical way to understand an important concept only once (.03 times per classroom observation).

After instruction, the teachers' scores on the checklist reached or exceeded the mastery level of 85% in 31 of the 32 observations. Teacher scores after instruction ranged from 35% to 100% (M = 94%). Teacher 6 failed to reach the mastery criterion the first time that she implemented the routine. She told the students that three concepts were important to understand, but she used the routine with only the third concept. Her scores during these three opportunities to use the routine were 3%, 6%, and 96%, respectively. Thus, her mean score for the class period was 35%. After consulting with the researchers, Teacher 6 earned scores above 95% during the remaining opportunities to use the routine.

After instruction in the use of the routine, the teachers coconstructed a total of 33 Concept Anchoring Tables with their students in 32 class sessions. (One teacher, Teacher 9, coconstructed two Concept Anchoring Tables in one class period with her students.) The teachers cued their students 30 times that the concept to be discussed was important and proceeded to implement the rest of the routine. On three occasions, teachers coconstructed Concept Anchoring Tables without presenting an importance cue. On only two occasions after introduction to the routine did a teacher present a cue that a concept was important to understand without presenting an analogy to help students understand the information. After the workshop, the teachers provided more than twice as many cues that information was important to understand as they did during baseline. Furthermore, they followed those cues with analogies

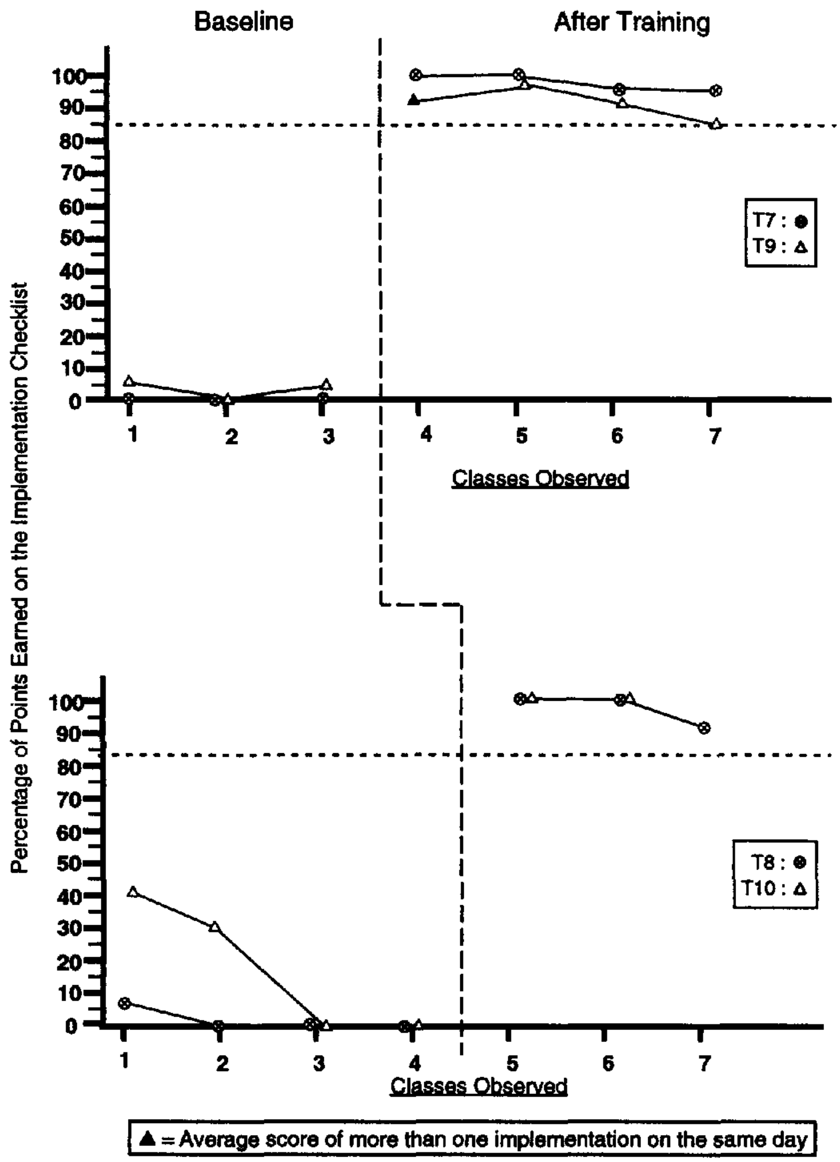

Figure 5. Percentage of points earned on the Implementation Checklist by Teachers (T) 7–10.

in 30 instances after training, as compared with one instance prior to training.

When interviews were conducted with the 10 teachers 8 years after the study was completed (1991), 8 of them reported that they continued to use the routine and the Concept Anchoring Tables they developed for the study when they taught the courses for which the tables were developed. A few of them indicated that they had made additional tables after the study was completed, but most indicated that they did not have the needed time to develop more.

Types of Analogies

The science teachers created 10 between-domains analogies and 7 within-domain analogies; the social studies teachers created 9 between-domains analogies and 7 within-domain analogies. The science teachers chose 8 concrete and 9 abstract known concepts, whereas the social studies teachers chose 5 concrete and 11 abstract known concepts. Finally, the science teachers created 4 structural analogies, 10 functional analogies, and 3 analogies that were both structural and functional. The social studies teachers created 4 structural analogies, 11 functional analogies, and 1 analogy that was both structural and functional. See Figure 6 for visual comparisons between the two types of teachers.

In general, the teachers selected concepts that were complex. For example, the science teachers selected concepts such as DNA, genetic counseling, covalent bonding, AIDS, inborn behavior, territoriality of animals, cellular respiration, biological molecules, the brain, and cellular respiration. Social studies teachers selected concepts such as the government of China, separate but equal schools, the relationships of territorial units in the British Isles,

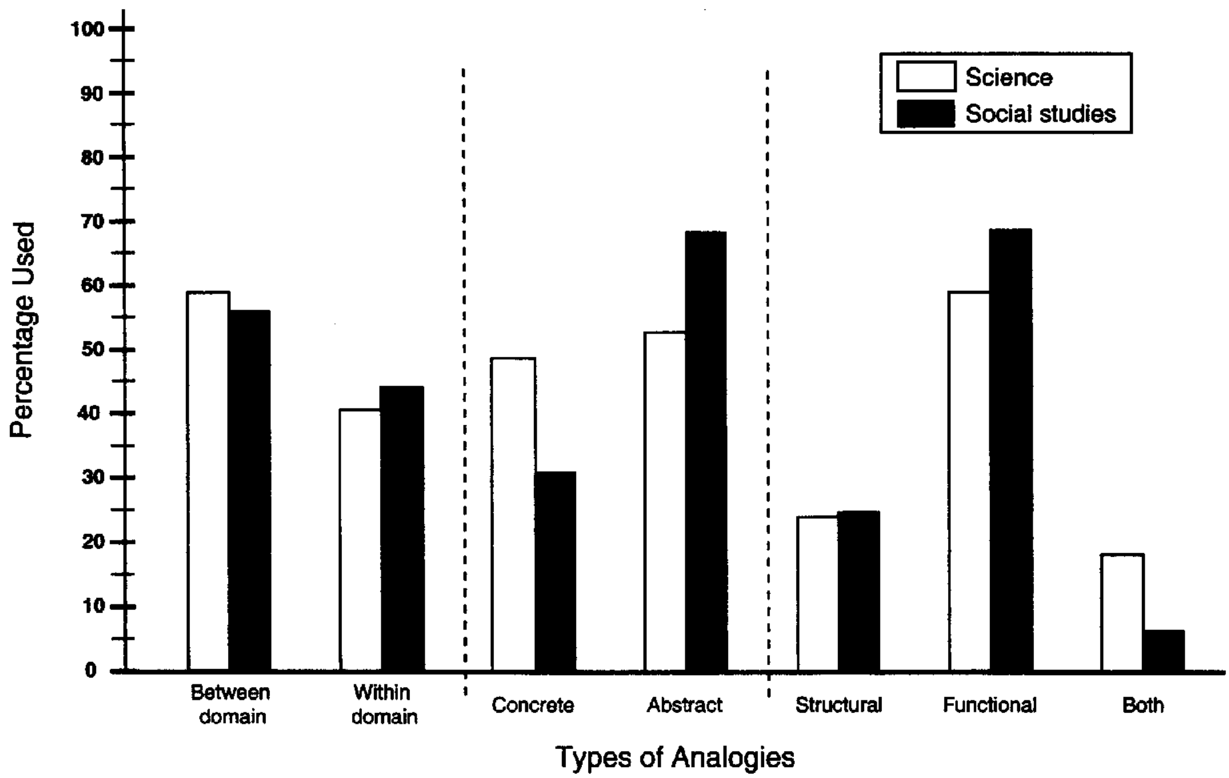

Figure 6. Percentages of different types of analogies developed by the teachers.

Australian aborigines, effects of the Free Soil Movement, and the history and government of India.

Satisfaction Ratings

The mean satisfaction ratings provided by the teachers on the Teacher Satisfaction Questionnaire indicated, on a 7-point scale, how satisfied the teachers were with factors related to the routine. On the scale, a *1* meant that teachers were completely dissatisfied, a *7* indicated that teachers were completely satisfied, and a *4* meant that teachers were neither satisfied nor dissatisfied. Results were as follows: flexibility of the routine, 5.9 (range = 4–7); ease of use, 5.5 (range = 3–7); acceptable preparation time, 5.3 (range = 3–7); that the table helped students with disabilities to learn facts, 5.6 (range = 3–7); that the table helped students without disabilities to learn facts, 5.7 (range = 4–7); that students with disabilities perceived the table to be useful, 4.8 (range = 2–7); that students without disabilities perceived the table to be useful, 5.1 (range = 3–7); that achievement for students with disabilities improved as a result of using the routine, 4.8 (range = 3–7); that achievement for students without disabilities improved as a result of using the routine, 5.0 (range = 4–7); that attention increased for students with disabilities, 5.4 (range = 3–7); that attention increased for students without disabilities, 5.2 (range = 3–7); that note-taking skills increased for students with disabilities, 5.1 (range = 4–7); that note-taking skills increased for students without disabilities, 4.8 (range = 1–6); that study time increased for students with disabilities, 3.8 (range = 1–6); that study time increased for students without disabilities, 4.0 (range = 1–6); that students with disabilities used the table to study for tests, 4.8 (range = 2–7); and that students without disabilities used the table to study for tests, 5.0 (range = 4–7). For the likelihood ratings, teachers indicated a high likelihood that they would continue to use the routine in their classes (M = 6.0, range = 4–7), that they would recommend the routine to other teachers (M = 5.7, range = 3–7), and that they would recommend it to others if inservice instruction were available (M = 5.9, range = 4–7).

Mean satisfaction ratings provided by the students were as follows: that the table helped them follow what the teacher was saying, 4.5 (range of class means = 3.3–5.4); that the table helped them take notes, 4.4 (range of class means = 3.6–6.1); that the table helped them to focus their attention on what was important in class, 4.5 (range of class means = 3.2–5.8); that the table helped them study for tests, 4.1 (range of class means = 2.7–5.9); that the table helped them do well on tests, 4.1 (range of class means = 2.8–5.7); how this new way of teaching compared with the traditional way, 4.3 (range of class means = 3.0–5.4); and that the table helped improve their grades, 3.7 (range of class means = 2.2–4.9).

In summary, teachers quickly coconstructed Concept Anchoring Tables in class with their students using the Concept Anchoring Routine at a high level of fidelity. All but 1 of the teachers exceeded the required mastery level the first time they used the routine in class. When they reached the mastery level, they maintained it. Both social studies and science teachers used all of the types of analogies. The social studies teachers used twice as many abstract analogies as concrete analogies; the science teachers used almost equal numbers of these two types. Otherwise, the two groups of teachers used about the same numbers of the different types of analogies. In general, the teachers were satisfied with the routine and indicated that they would continue to use it and recommend it to others. The students were less satisfied than their teachers, but, with one exception, their mean scores on the items did not fall in the dissatisfied range. There was a wide variation among the satisfaction of different classes of students, indicating that the different teachers may have implemented the steps of the routine in different ways or with different levels of enthusiasm or that some teachers' Concept Anchoring Tables were more helpful than others'. Another possibility for the discrepancy in student satisfaction was that the students might not have been benefiting from the use of the routine in terms of improved performance when it was used in actual classrooms with actual course content.

Study 3: Effects of Routine Use in an Actual Class

Study 3 was conducted to determine the effects of the Concept Anchoring Routine on student knowledge of actual course content when the routine was implemented by a general education teacher.

Method

Participants and Setting

One of the 10 teachers who participated in Study 2 (Teacher 8) also participated in Study 3. She taught seventh-grade, general education life science classes that contained a diversity of students. The teacher had used the Concept Anchoring Table and Concept Anchoring Routine in her classes, as described in Study 2. Eighteen students in one of her classes that had participated in Study 2 also participated in Study 3. The setting was the teacher's regularly assigned classroom.

Measurement System

Four parallel, equivalent forms of a nine-item test were designed to measure student recall of information related to four targeted concepts. Through the use of nine standard questions, each instrument assessed a student's ability to provide the following information about the specified concept: words associated with the concept, a definition of the concept, a description of the concept, characteristics of the concept, something similar to the concept, something different from the concept, the name of the larger group to which the concept belongs, a summary of knowledge about the concept, and an explanation of how the concept affects the student's life or the world. Each item was an open-ended question eliciting the specified type of information (e.g., What are characteristics of alveoli?). To aid scorers, a correct answer was specified in writing for each question on each test. Each student answer was scored as correct or incorrect on the basis of the answer key, and each correct answer was awarded 10 points, for a total of 90 possible points on the test. No partial credit was awarded. A percentage correct score was calculated for each student for each test.

Interscorer reliability on student tests was determined by having two scorers independently score 15% of the tests. The scorers were blind to the purpose and procedures of the study. Their scores were compared item by item, and an agreement was scored if both scorers indicated that an item was correct or if both scorers indicated it was incorrect. The percentage of agreement was calculated as described earlier; it was 99% (98 agreements within 99 opportunities for agreement). The percentage of agreement on individual tests ranged from 89% to 100%.

Procedures

Teacher 8 chose four concepts for this study: epiglottis, pancreas, alveoli, and esophagus. Three science experts were asked to review the concepts and the information associated with them and then to pair the concepts according to the complexity of the concept and the associated information. They independently paired epiglottis with pancreas and alveoli with esophagus, judging that the paired concepts were equivalent in complexity. Then, two of the concepts, epiglottis and alveoli, were randomly chosen to be taught with a Concept Anchoring Table and the Concept Anchoring Routine. The two others, esophagus and pancreas, were assigned to be taught with the traditional lecture–discussion format. The teacher then designed a short, 15-min lesson about each of the concepts. All four lessons had the same number of characteristics associated with the targeted concept. Next, the teacher gave the lessons in class in sequence with one day between the lessons. The teacher wrote the characteristics associated with each of the concepts on the board during each lesson and prompted the students to take notes. The only difference between the lessons was that the teacher gave the students a blank Concept Anchoring Table at the beginning of the enhanced lessons and wrote characteristics associated with the known concept on the board when an analogy was used.

Experimental Design

An ABAB reversal design (Baer et al., 1968) was used. Each concept was taught in a separate lesson. Epiglottis was taught first and enhanced with the Concept Anchoring Routine and table, pancreas was taught second with the traditional lecture–discussion format, alveoli was taught third and enhanced with concept anchoring, and esophagus was taught last with the traditional lecture–discussion format. The test about the concept was administered by the teacher on the day following the instruction on that concept. For each of the concepts taught, the mean percentage of points earned on the test by the class was used to assess student understanding and memory of information related to that concept.

Results

After epiglottis was taught using the routine, the students earned an average test score of 83% (range = 44–100%). After they were taught the concept of pancreas through traditional methods, they earned an average test score of 27% (range = 0–89%) on the test. After alveoli was taught using the routine, the students earned an average test score of 70% (range = 0–100%). When esophagus was taught using traditional methods, students earned an average score of 42% (range = 0–78%).

The students' mean test performance on the two enhanced concepts was compared with their mean test performance on the two nonenhanced concepts using t tests. A significant difference was found in favor of the concepts associated with the routine, $t(34) = 9.11$, $p < .000$. When comparisons were made on each pair of concepts, significant differences were also found in favor of the concept that had been associated with the routine in each pair (for epiglottis vs. pancreas, $t(34) = 7.33$, $p < .000$; for alveoli vs. esophagus, $t(34) = 3.35$, $p < .002$).

Conclusions

The results of Study 3 provide a preliminary indication that the use of the routine aids student learning about concepts taught in an actual class. Student performance associated with the routine translated into substantially more students earning passing grades as well as substantially more students earning average and above-average grades. Such a difference is critical if students are to feel good about themselves and their learning and if at-risk students are to benefit from instruction offered in the general education classroom. Furthermore, teachers are likely to continue using new educational techniques if gains are apparent on measures they already consider important, such as classroom tests (Fullan, 1982).

Discussion

The results of this project indicate that the use of analogical instruction (in the form of the Concept Anchoring Routine and table) can enhance student performance, with regard to understanding and remembering secondary subject-matter content, to such an extent that many more students not only pass tests but perform at commonly acceptable levels on tests than if they receive traditional instruction. In addition, analogical instruction appears to be powerful enough to enhance the performance of students who are at risk for failure as well as other students in diverse classes. Furthermore, the results indicate that analogical instruction not only produces positive results in laboratory-type classroom arrangements (i.e., with a researcher doing the instruction as in Study 1) but also in an actual secondary class taught by a subject-area teacher under typical school conditions (as in Study 3). The results also show that teachers can quickly learn to create Concept Anchoring Tables and use the Concept Anchoring Routine and that they are satisfied with the routine.

The fact that improvements in performance were realized by all four types of students who participated in Study 1 is important. Because many inclusive secondary classrooms contain a majority of students who fall in the NA and HA range, the impact of any educational innovation on the performance of these students is critical. Indeed, Schumaker et al. (1991) have reported that in order for teachers to continue using an instructional innovation, they must be convinced that it benefits the largest group of students in the class (generally, the NA and HA students). Therefore, the fact that the NA and HA students' performance was positively affected by the routine indicates that it has promise for being maintained over time in classrooms. In addition, the fact that the performance of LA and LD students was positively affected is important because these are the very students who need the most assistance. Indeed, teachers of academically diverse classes are expected, more and more, to arrange instruction so that all students can successfully reach certain standards. In both Studies 1 and 3, use of the routine was associated not only with many more students earning passing scores on the tests but also with more students earning average and above-average scores.

Nevertheless, several issues remain. First, in both Studies 1 and 3, although their performance improved, some students continued to earn failing grades when the routine was used. The characteristics of students who continue to fail and reasons for their failure (e.g., high rates of absenteeism, not using the Concept Anchoring Table to study for a test, not participating during the construction of the table, not taking adequate notes on the blank form of the table) need to be identified. In addition, elements that can be added to the routine that might help them perform at passing levels need to be considered and explored. For example, they might need instruction on how to study the Concept Anchoring Table before a test, or they might need additional support as they study for tests.

Second, mean student satisfaction ratings were not within the "satisfied" range (5.0–7.0 on the 7-point scale) but were in the "neutral" range and varied greatly across classes associated with different teachers. Several possible reasons for the lower satisfaction ratings and this variation need to be explored. First, the routine was introduced in the classes well after the school year was underway, and the students were accustomed to the ways in which their teachers had been teaching. Informal discussions with the teachers indicated that students often embrace routines introduced at the beginning of the school year and resent new routines that are introduced later in the year. Perhaps if the routine had been introduced at the beginning of the school year, the students might have been more positive. Third, the teachers may have been using the routine in qualitatively different ways. Even though their scores on the checklist were similar, their enthusiasm for the routine might have differed, or the quality or types of their analogies may have differed. More specific ways of measuring teacher performance need to be developed, and the performances of teachers who receive higher ratings and lower ratings need to be compared. Fourth, the routine may not have included elements that promote student "buy-in." Elements that might be added include rationales regarding the value of the use of analogies, a close tie between the information taught with the routine and assessments, reviews prior to tests highlighting information taught with the routine, and feedback regarding the efficacy of the routine related to gains in understanding and performance on tests. If students are truly to be partners in a learning community (Bulgren et al., 1994a), they must feel that they understand, participate in, and value the routines used in a class.

In addition, the differential effects of different types of analogies need to be determined. Study 1 provides some preliminary evidence that different types of analogies might yield different results. For commensalism, an abstract, functional analogy was developed; for pyramid of numbers, a concrete, structural analogy was developed. Although the HA students scored similarly on both concepts (*M*s = 96% and 94%, respectively), the NA students, on the average, performed a letter grade higher on the concept associated with the concrete analogy (M = 92%) than on the one associated with the abstract analogy (M = 84%). In contrast, the LA students scored slightly higher on the concept associated with the abstract analogy (80% vs. 73%). Students with LD scored higher on the concept taught with the concrete analogy (M = 69%) than on the concept taught with the abstract analogy (M = 55%). As a result, 77% of the students with LD earned scores of 60% or above on pyramid of numbers items, whereas only 36% of the students with LD earned scores at or above the 60% level on commensalism items. The difference in performance on the two concepts was approximately twice as large for students with LD as the differences found between types of analogies for LA and NA students, and this finding was not surprising given the fact that the majority of secondary students with LD are reasoning at the concrete level (Skrtic, 1980). These findings may provide some support for recommendations that analogies be drawn from the concrete realm as much as possible (Keane, 1987).

However, there are other possible explanations for the differential effects of the different analogies. One is that the students may not have been as familiar with a story of a lemonade stand as they were with the levels of military structure. As a result, the story associated with the lemonade stand might have created an additional burden for students instead of enhancing their understanding of the concept. Alternatively, the commensalism concept might have been more complex than the pyramid of numbers concept.

Clearly, additional research is needed that explores the different dimensions associated with analogies. Future research might also explore the types of analogies teachers and students find to be most effective and least effective and why.

Overall, the results of the three studies must be tempered by their limitations. First, the investigation of the effects of the routine on student performance (Studies 1 and 3) was limited to science content. Although Study 2 was conducted in both science and social studies classes, no data were gathered on students' understanding of social studies content when the routine was associated with that content. In addition, teacher data are not available on the use of analogies in other subject areas. Studies are clearly needed in other content and vocational areas.

Second, the use of intact classrooms in Study 1 resulted in small numbers and unbalanced numbers of students in some of the subgroups, which may have affected the statistical comparisons. Studies that include larger groups of students would provide more information regarding which types of analogies are most useful with each group.

Third, future studies in actual classrooms in a variety of subject areas under typical school conditions are needed to determine student effects as well as to determine whether teachers will continue to construct additional analogies and use the routine over time. The baseline condition in Study 2 indicated that the participating teachers used few analogies during instruction. Thus, the use of analogies does not appear to be a natural part of instruction and may require considerable support to sustain.

Fourth, the teachers who participated were volunteers. Whether the same kinds of results can be achieved with teachers who do not voluntarily learn the routine for the purpose of participating in a research study is unclear. Also, the school districts that participated are suburban districts; whether the same kinds of results can be achieved in other types of school settings is not known. Additional research is needed in secondary general education classrooms in rural and urban school districts and in other subject areas to replicate and extend these results. As a corollary, the different types of analogies developed by teachers across subject areas, in different types of schools, and at different grade levels need to be identified and compiled.

Fifth, the length of time during which students received the enhanced instruction was relatively short, and the number of times was limited. There is a need to explore effects of the routine on student learning and satisfaction over a longer period of time (e.g., a whole school year) in academically diverse classes.

Sixth, although many of the teachers continue to use the routine, they indicated that they have continued to use the analogies they developed during the study and have rarely developed new ones because of time constraints. Thus, future research needs to explore ways of providing teachers with the time they need to develop new instructional tools. Alternatively, curriculum development efforts need to focus on providing analogies for commonly taught complex concepts within standard courses.

Finally, this project focused on the effects of the Concept Anchoring Routine in isolation. Additional research is needed to investigate the effects of the Concept Anchoring Routine combined with other Content Enhancement Routines such as the Concept Mastery Routine (Bulgren, Deshler, & Schumaker, 1993), the Concept Comparison Routine (Bulgren, Lenz, Deshler, & Schumaker, 1995), the Lesson Organizer Routine (Lenz, Marrs, Schumaker, & Deshler, 1993), the Recall Enhancement Routine (Schumaker, Bulgren, Deshler, & Lenz, 1998), and the Unit Organizer Routine (Lenz, with Bulgren, Schumaker, Deshler & Boudah, 1994) as well as other instructional routines developed to enhance learning in academically diverse classes (e.g., Carnine & Shinn, 1994; Johnson & Johnson, 1989; Slavin, 1996). Such research would ideally incorporate many of the elements included in this series of studies, including carefully structured stages of development that lead from research to practice and different types of research methodologies which, taken together, provide validation for research-based instruction that can respond to all of the increasingly complex challenges facing students and teachers in general content classrooms.

The results of this investigation and other studies related to the other Content Enhancement Routines (e.g., Bulgren et al., 1988; Bulgren, Schumaker, & Deshler, 1994b) have implications for preservice teacher training and ongoing staff development. The growing body of literature in this area is showing that changes in instruction can have a socially, as well as a statistically, significant impact on student learning in secondary general education classes in which diverse groups of learners are enrolled (i.e., these changes enable failing students to pass, and barely passing students to succeed). Such findings lend support for emphasizing change at the level of instruction in secondary schools, and such change will germinate from preservice and inservice training for secondary teachers in validated instructional methods. Teachers need to be equipped to meet the pressures they face with regard to teaching students more and more complex information and higher order thinking processes. Instruction in routines such as the Concept Anchoring Routine can be one answer to such pressures.

References

Andrews, A. C. (1987, September–October). The analogy theme in geography. *Journal of Geography, 86,* 194–197.

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning.* New York: Grune & Stratton.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1,* 91–97.

Bean, T. W., Singer, H., & Cowan, S. (1985, December). Analogical study guides: Improving comprehension in science. *Journal of Reading, 29,* 246–250.

Berliner, D. C. (1987). But do they understand? In V. Richardson-Koehler (Ed.), *Educators' handbook: A research perspective* (pp. 259–293). New York: Longman.

Brown, A. L., Campione, J. C., & Day, J. D. (1981, February). Learning to learn: On training students to learn from texts. *Educational Researcher, 10,* 14–21.

Bulgren, J. A., Deshler, D. D., & Schumaker, J. B. (1993). *The Concept Mastery Routine.* Lawrence, KS: Edge Enterprises.

Bulgren, J. A., & Lenz, B. K. (1996). Strategic instruction in the content areas. In D. D. Deshler, E. S. Ellis, and B. K. Lenz (Eds.), *Teaching adolescents with learning disabilities: Strategies and methods* (2nd ed., pp. 409–473). Denver, CO: Love Publishing.

Bulgren, J. A., Lenz, B. K., Deshler, D. D., & Schumaker, J. B. (1995). *The Concept Comparison Routine.* Lawrence, KS: Edge Enterprises.

Bulgren, J. A., Lenz, B. K., Scanlon, D., & Clark, C. (1995, April). *Participatory research: Varied methodologies, issues, and findings in development and implementation of research partnerships involving teachers and researchers.* Paper presented at the American Educational Research Association Conference, San Francisco.

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,* 3–17.

Bulgren, J. A., Schumaker, J. B., & Deshler, D. D. (1994a). *The Concept Anchoring Routine.* Lawrence, KS: Edge Enterprises.

Bulgren, J. A., Schumaker, J. B., & Deshler, D. D. (1994b). The effects of a recall enhancement routine on the test performance of students with and without learning disabilities. *Learning Disabilities Research and Practice, 9,* 2–11.

Carnine, D., & Shinn, M. (Eds.). (1994). Diverse learners and prevailing, emerging, and research-based educational approaches and their tools. *School Psychology Review, 23,* 341–471.

Christian, M. C. (1990, Spring/Summer). HERD. *NEACT,* 9–11.

Connell, P. J. (1987). Teaching language rules as solutions to language problems: A baseball analogy. *Language, Speech, and Hearing Services in Schools, 18,* 194–205.

Curtis, R. V. (1988). When is a science analogy like a social studies analogy? A comparison of text analogies across two disciplines. *Instructional Science, 17,* 169–177.

Fullan, M. (1982). *The meaning of educational change.* New York: Teachers College Press.

Gagne, R. M. (1965). *The conditions of learning.* New York: Holt, Rinehart & Winston.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7,* 155–170.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10,* 277–300.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306–355.

Goldman, S. R., Pellegrino, J. W., Parseghian, P., & Sallis, R. (1982). Developmental and individual differences in verbal analogical reasoning. *Child Development, 53,* 550–559.

Holyoak, K. J. (1984). Analogical thinking and human intelligence. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 199–230). Hillsdale, NJ: Erlbaum.

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research.* Edina, MN: Interaction Book Co.

Keane, M. (1987). Cognitive theory of analogy. In J. T. E. Richardson, M. W. Eysenck, & D. W. Piper (Eds.), *Student learning: Research in education and cognitive psychology* (pp. 67–72). Philadelphia: Open University Press.

Lenz, B. K., Alley, G. R., & Schumaker, J. B. (1987). Activating the inactive learner: Advance organizers in the secondary content classroom. *Learning Disability Quarterly, 10,* 53–68.

Lenz, B. K., Bulgren, J. A., & Hudson, P. (1990). Content enhancement: A model for promoting the acquisition of content by individuals with learning disabilities. In T. E. Scruggs & B. L. Y. Wong (Eds.), *Intervention research in learning disabilities* (pp. 122–165). New York: Springer-Verlag.

Lenz, B. K. (with Bulgren, J. A., Schumaker, J. B., Deshler, D. D., & Boudah, D. A.). (1994). *The Unit Organizer Routine.* Lawrence, KS: Edge Enterprises.

Lenz, B. K., Marrs, R. W., Schumaker, J. B., & Deshler, D. D. (1993). *The content enhancement series: The Lesson Organizer Routine.* Lawrence, KS: Edge Enterprises.

Lenz, B. K., Schumaker, J. B., & Deshler, D. D. (1991, March). Planning in the face of academic diversity: Whose questions should we be answering? Paper presented at the American Educational Research Association Conference, Chicago.

Markman, E. M. (1985). Comprehension monitoring of developmental and educational issues. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Research and open questions* (Vol. 2, pp. 275–291). Hillsdale, NJ: Erlbaum.

Mayer, R. E. (1987). *Educational psychology: A cognitive approach.* Boston: Little, Brown.

McGonigal, E. (1988). Correlative thinking: Writing analogies about literature. *English Journal, 77,* 66–67.

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

Pressley, M. (with McCormick, C. B.). (1995). *Cognition, teaching, & assessment.* New York: HarperCollins.

Putnam, M. L., Deshler, D. D., & Schumaker, J. B. (1992). The investigation of setting demands: A missing link in learning strategy instruction. In L. J. Meltzer (Ed.), *Strategy assessment and instruction for students with learning disabilities: From theory to practice* (pp. 325–351). Austin, TX: Pro-ed.

Raphael, T. E., & Gavalek, J. R. (1984). Question-related activities and their relationship to reading comprehension: Some instructional implications. In G. G. Duffy, L. R. Roehler, & L. J. Mason (Eds.), *Comprehension instruction: Perspectives and suggestions* (pp. 234–250). New York: Longman.

Rumelhart, D. E., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 95–135). Hillsdale, NJ: Erlbaum.

Schumaker, J. B., Bulgren, J. A., Deshler, D. D., & Lenz, B. K. (1998). *The Recall Enhancement Routine.* Lawrence, KS: Edge Enterprises.

Schumaker, J. B., Deshler, D. D., & McKnight, P. (1991). Teaching routines for content areas at the secondary level. In G. Stover, M. R. Shinn, & H. M. Walker (Eds.), *Interventions for achievement and behavior problems* (pp. 473–494). Washington, DC: National Association of School Psychologists.

Sheard, C., & Readence, J. E. (1988). An investigation of the inference and mapping processes of the componential theory of analogical reasoning. *Journal of Educational Research, 8,* 347–353.

Skrtic, T. M. (1980). *Formal reasoning abilities of learning disabled adolescents* (Research Brief No. 7). Lawrence: University of Kansas, Institute for Research in Learning Disabilities.

Slavin, R. E. (1996). Research for the future: Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21,* 43–69.

Sternberg, R. J., & Ketron, J. L. (1982). Selection and implementation of strategies in reasoning by analogy. *Journal of Educational Psychology, 74,* 399–413.

Vosniadou, S., & Ortony, A. (1989). Similarity and analogical reasoning: A synthesis. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 1–17). Cambridge, England: Cambridge University Press.

White, C. S., & Alexander, P. A. (1984). Teaching analogical reasoning processes. *Reading World, 24,* 38–42.

Wong, B. Y. L. (1985). Potential means of enhancing content skills acquisition in learning disabled adolescents. *Focus on Exceptional Children, 17,* 1–8.

Wragg, P. H., & Allen, R. F. (1983). Developing creativity in social studies III: Generating analogies. *Georgia Social Science Journal, 14,* 27–33.

Received March 23, 1999
Revision received October 15, 1999
Accepted October 15, 1999 ■