Learning Disabilities Research & Practice, 9(1), 2–11

The Effects of a Recall Enhancement Routine on the Test Performance of Secondary Students With and Without Learning Disabilities

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

The purpose of this study was to evaluate the effects of presenting mnemonic devices in conjunction with content information on the recall performance of students with and without learning disabilities (LDs). These devices were presented using a standardized set of procedures called the Recall Enhancement Routine. Students were randomly assigned to an experimental or a control group. All students received a lecture containing the same content. During a review portion of the lesson, the Recall Enhancement Routine was used for students in the experimental group in conjunction with some of the information in the lecture. For students in the control group, repetitious review was presented in conjunction with the same information. The study took place in the mainstream classrooms of secondary content teachers who allowed their classrooms to be used in the study. Information and testing procedures validated by the teachers as typical for their classes were developed. Results showed that students with and without LDs within the experimental group recalled significantly more of the reviewed information than students in the control group. In addition, compared to students in the control group, substantially more students in the experimental group scored within passing range on the reviewed items. Therefore, the Recall Enhancement Routine has potential for use by regular classroom teachers to facilitate the inclusion of students with LDs within academically diverse classes of students.

Increasingly, students with learning disabilities (LDs) are being included in mainstream classes where they must confront and overcome curricular demands that are characterized by large volumes of information, complexity, and abstractness. This is especially true in secondary settings (Deshler & Schumaker, 1988), where demands have been exacerbated by recent calls for raising performance standards and increasing the amount of content covered in core classes (U.S. Department of Education, 1991). In addition to possibly increasing the pressures on students with LDs, inclusionary placements also present major challenges for regular classroom teachers (Schumaker & Deshler, 1988). Indeed, today's teachers are expected to emphasize student comprehension of major concepts, trends, and issues, as well as to teach large amounts of factual information.

In spite of the objections of some (e.g., Edgar, 1993; Poplin, 1992), educators increasingly expect students to acquire factual information, particularly at the secondary level (Schumaker, Deshler, & McKnight, 1991). Putnam, Deshler, and Schumaker (1992) documented a similar finding in a survey of setting demands that 7- and 10th-grade students are expected to meet based on teacher- and publisher-made tests in the regular classroom. For example, in determining demands on students to respond to test questions, these researchers found that the largest percentage of test questions required students to recall specific facts at the knowledge level of Bloom's (1956) taxonomy of educational objectives. On the average, almost 50% of the questions were basic recognition questions (i.e., multiple-choice, matching, and true–false), and an additional 10% were recall questions (e.g., those requiring students to retrieve the correct answer from memory and write it in a brief form).

Scruggs and Mastropieri (1990b) argued that instructional procedures that promote acquisition of factual knowledge (e.g., the use of mnemonic devices) may facilitate comprehension and conceptual understanding. This occurs because mnemonic systems tend to enhance the concreteness and meaningfulness of the targeted information and, in turn, its comprehension. Several studies involving students with LDs have investigated the effectiveness of using a variety of mnemonic devices (e.g., keywords, visual representations, paired associates, etc.) to facilitate student performance. These investigations have been conducted either in special education settings or under laboratory-type conditions. For example, Mastropieri, Scruggs, McLoone, and Levin (1985) demonstrated that students with LDs who were individually

Requests for reprints should be sent to Janis A. Bulgren, University of Kansas, Center for Research on Learning, Robert Dole Human Development Center, Room 3061, Lawrence, KS 66045.

provided visual representations of keywords representing the information to be remembered outperformed both students who were instructed with direct instruction and students who were allowed to study the information on their own. Similarly, Condus, Marshall, and Miller (1986) found keyword–image mnemonics to enhance the performance of 12-year-old LD students who were taught in resource room settings.

Using computer-presented stimuli, Ferro and Pressley (1991) noted that the performance of both students with and without LDs who constructed visual images to link paired associates was significantly better than the performance of similar students who simply rehearsed the information. Scruggs and Mastropieri (1989) and Mastropieri and Scruggs (1988) implemented mnemonic instruction to teach American history facts in high school self-contained classrooms. The students who were instructed through the use of mnemonic devices in these settings scored much higher on tests than did students who were instructed traditionally. King-Sears, Mercer, and Sindelar (1992) tested the use of mnemonic devices in special education science classrooms for students with LDs and emotional disturbances. They also found significantly better short-term retention of information on matching test items when teacher-prepared and student-prepared keywords were learned than when traditional instructional methods were used. Finally, Bulgren, Hock, Schumaker, and Deshler (1993) found that students with LDs were able to master a strategy involving the use of various types of mnemonic devices when taught by a special educator in a small-group instructional setting. Students were able to use the strategy generatively when learning and memorizing factual information in passages derived from secondary textbooks. Further, the students' use of the strategy was associated with substantial improvement in recall performance.

Collectively, these studies demonstrate that students with LDs can learn mnemonic strategies and successfully use such devices to improve their performance in both special education and experimentally controlled settings. Scruggs and Mastropieri (1990a) also pointed out the importance of determining the applicability of mnemonic strategies within the context of real classroom settings. Their research is an excellent example of how theory-driven, laboratory-based research can be extended to address the realities of the classroom. They also emphasized, through their long line of research on mnemonic strategies, the need to vary experimental designs, materials, and procedures as a function of the setting and other experimental conditions. In short, they argued that research must ultimately be conducted under instructional conditions that are comparable to those that exist in real classrooms, resulting in an instructional dynamic that is authentic in nature.

The purpose of this study was to shed light on how students with LDs could be taught to better cope with the realities of secondary school curriculum demands. More specifically, we attempted to reflect the challenge facing the regular classroom teacher of having to present content information so that students with LDs can understand and remember the critical elements of the lesson while maintaining the integrity of the content. Therefore, we extended previous research by determining the effects of presenting three types of mnemonic devices (acronyms, visual images, and keywords) in conjunction with factual content on the recall performance of students enrolled in regular secondary classrooms, including a broad diversity of students (i.e., students with and without LDs).

The instructional routine designed for this study was based on a programmatic line of research conducted by investigators at the University of Kansas Center for Research on Learning (KU–CRL) (e.g., Lenz & Bulgren, in press; Lenz, Bulgren, & Hudson, 1990; Schumaker et al., 1991), in which content enhancement routines have been used to facilitate the teaching of scientific or cultural knowledge to heterogeneous groups of students in regular classes. Other research with LD students, conducted by the KU–CRL staff, suggests that these students' performance can be improved if teachers modify the way content information is organized and presented to the whole class. For example, published research studies have focused on the introduction of a lesson (Lenz, Alley, & Schumaker, 1987) and the teaching of concepts (Bulgren, Schumaker, & Deshler, 1988) within secondary mainstream class formats. These studies were designed to validate methods to help teachers increase student comprehension of large amounts of content information in regular education classrooms.

The research reported in this study acknowledges the need to help students recall in addition to comprehend and manipulate information. For the purposes of this study, the following criteria have been specified as central to effective content delivery to academically diverse student populations: (a) both group and individual learning needs must be met; (b) the integrity of the content must be maintained; (c) critical features of the content must be selected, organized, manipulated, and enhanced in a manner that promotes effective and efficient information processing; and (d) the content must be delivered in such a way that the learning of all students is facilitated and enriched. These considerations are based on research indicating that mainstream teachers are reluctant to use procedures that target only a few students (Lenz, Schumaker, & Deshler, 1991).

METHOD

Subjects

Participants included 41 students in the seventh and eighth grade at a junior high school in a suburban Midwestern school district. The students were recruited from two social studies classes that were team taught by a special education teacher and a social studies teacher. For the selection process, students from both classes who volunteered for the study were stratified by grade level (seventh or eighth) and condition (i.e., LD or nonlearning disabled [NLD]). Half of the students in each stratified group were randomly selected to comprise the experimental group; the remaining students comprised the control group. Specifically, participants consisted of 18 LD students (9 in the seventh grade and 9 in the eighth grade) and 23 NLD students (11 in the seventh grade and 12 in the eighth grade). The experimental group was composed of 9 LD students and 11 NLD students; the control group was composed of 9 LD students and 12 NLD students.

Students with LDs had been classified as such by school personnel following district and state guidelines for identify-

TABLE 1
Description of Subjects

	Experimental Group	*Control Group*
Numbers		
Male	5	7
Female	4	2
Total	9	9
Age		
Mean	14 years, 2 months	14 years, 3 months
Range	3 years, 2 months	1 year, 10 months
Race/Ethnicity		
Anglo	8	9
Hispanic	–	–
African American	–	–
Native American	1	–
Asian American	–	–
Other	–	–
SES		
High/middle	6	8
Low	3	1
Grade level		
Seventh	4	5
Eighth	5	4
Intelligence (a)		
Mean SS	96.11	97.11
SD	12.56	12.36
Range	35.00	37.00
Overall academic achievement (b)		
Mean SS	89.89	95.44
SD	10.08	18.26
Range	30.00	60.00
Specific academic achievement (c)		
Mean SS	89.89	89.11
SD	6.41	7.46
Range	18.00	20.00
Location		
Geographic region	Midwest	Midwest
Locale	Suburban	Suburban
a. Name(s) of test(s) used:	WISC–R Full Scale	
b. Name(s) of test(s) used:	Woodcock–Johnson Knowledge	
c. Name(s) of test(s) used:	Woodcock–Johnson Reading	

ing LD students. See Table 1 for demographic data on the LD students. The experimental group included four LD students in the seventh grade and five in the eighth grade; the control group included five LD students in the seventh grade and four in the eighth grade. The experimental group included five male LD students and four female LD students; the control group included seven male LD students and two female LD students.

The NLD students were students in the two participating social studies classes who had never been classified as qualifying for special services. The experimental group consisted of five males and six females; the control group consisted of seven males and five females. In terms of grade level, the experimental group included four NLD students in the seventh grade and seven in eighth grade, whereas the control group included six NLD students in the seventh grade and six in the eighth grade.

Settings

The settings for this study were the two mainstream social studies classrooms in a junior high school serving Grades 7 to 9, with a total enrollment of 655 students. These were typical classrooms with desks, chairs, and chalkboards.

The Recall Enhancement Routine

The Recall Enhancement Routine, which has been socially validated by teachers and shown to be usable in mainstream, secondary content classes in a companion study (Bulgren, Deshler, & Schumaker, 1993), is a teaching routine designed to be incorporated into normal mainstream instruction for the purpose of enhancing students' retention of targeted information by associating mnemonic devices with that information. The routine can be used at any time during a lesson; however, in our study, it was incorporated into the closing part of the lesson during a review of the content that had been presented during the lesson.

At the beginning of this review, the students were told that the information that had been presented would be reviewed in a way that would enhance their recall of the information. In essence, the instructor told the students that she had designed some memory devices that might help them remember the information. As each item of targeted information was reviewed, the Recall Enhancement Routine was implemented following this sequence: (a) the students were verbally cued that the information that was to follow was important; (b) the students were verbally cued to take notes about the information and the device for remembering it; (c) the type of mnemonic device to be used to remember the information was named (e.g., the students were told that they would be using a mental image to remember the information); (d) the mnemonic device that had been specifically designed for the information was presented in conjunction with that information, and both the device and the information were written or sketched on the board; and (e) the mnemonic device was reviewed at the end of the review period. (See Table 2 for an example script for the steps of the routine.)

The Content Lesson

A topic, history of American journalism, was selected as the lesson to be presented. This topic was selected because (a) a lesson could be designed about it that contained information of the type students in a social studies content classroom might be expected to remember, (b) the teachers of the two social studies classes agreed that the lesson contained valid content to be presented in their courses, and (c) the teachers concurred that the topic had not been covered in their courses and predicted that their students would have limited prior knowledge about it.

As the lesson was prepared, factual information judged important for the students to remember was selected (e.g., "William Randolph Hearst covered the Spanish-American War in Cuba" and "The code of ethics for newspaper editors

TABLE 2
Example Script for the Memory Enhancement Routine

Importance cued	"You'll need to remember the elements of the **code of ethics** for newspaper editors."
Notetaking cued	"Be sure to get the memory device for remembering this information and this information in your notes."
Mnemonic devices named	"To remember the elements of the newspaper editors' code of ethics, we'll use the acronym 'FAIR.' " (Teacher writes **Code of Ethics** on the board and then writes the word *FAIR* on the board vertically.)
Mnemonic device presented and linked to the information	" 'F' stands for **'fair'**; 'A' stands for **'accurate'**; 'I' stands for **'impartial'**; and 'R' stands for **'responsible.'** " (Teacher writes the four elements on the board next to their corresponding letters.) "The word 'FAIR' will help you remember that the code of ethics for newspaper editors requires them to be fair, accurate, impartial, and responsible. Use this sentence to remember the acronym 'FAIR': 'The code of ethics helps editors be FAIR.' "
Mnemonic device reviewed (at the end of the review period)	"What device will you use to remember the elements of the code of ethics for editors?" (Elicited from the students: "In the acronym 'FAIR,' 'F' stands for 'fair'; 'A' stands for 'accurate'; 'I' stands for 'impartial'; and 'R' stands for 'responsible.' ")

demanded that members of the profession be fair, accurate, impartial, and responsible") and woven into a set of lecture notes. (These sets of important information will be referred to hereafter as *facts*.) The lecture was designed to be 45-min long, containing a total of 40 targeted facts.

After teacher notes for the lecture had been prepared, the 40 facts were separated into three categories by four judges[1] according to the mnemonic device that was deemed to be most appropriate. The three types of mnemonic devices had been identified in another study by regular content teachers as the devices they most commonly used with their students (Bulgren, Deshler, & Schumaker, 1993). The three types of devices were acronyms, visual images, and keywords. The four judges concurred on the categorization of all of the facts.

Next, from the list of 40 facts, a test was designed (see the Measurement section for details) and administered to a group of students in another junior high school in the same school district. An index of difficulty was computed for each item on the test, corresponding to each of the targeted facts. The facts were then paired according to their index of difficulty. Subsequently, one fact in each pair was randomly selected to be reviewed at the end of the lesson (these will be referred to hereafter as *reviewed facts*); the other fact in each pair was not reviewed at the end of the lesson (these will be referred to hereafter to as *nonreviewed facts*). Within the 21 reviewed facts, 7 could be associated with images, 10 could be associated with acronyms, and 4 could be associated with keywords. Within the 19 nonreviewed facts, 8 facts could be associated with images, 5 could be associated with acronyms, and 6 could be associated with keywords. Once the lists had been finalized, a mnemonic device was designed for each of the reviewed facts according to the appropriate category. (See Table 3 for examples.)

Two lecture/discussion scripts were designed: one for the experimental group and one for the control group. Both scripts contained the same presentation of logically ordered information on the selected topic. (This portion of the lesson will be referred to hereafter as the *lecture portion*.) At the beginning of both scripts, statements verbally cued the students about the importance of the information and prompted the students to take notes. Imbedded in both scripts were statements to be made by the teacher about each of the 40 targeted facts, and additional information was provided to relate all the facts.

For example, when presenting information about William Randolph Hearst and the war in Cuba, the script read, in part:

> Another famous reporter was William Randolph **Hearst** who also owned many papers. He covered the Spanish-American War in **Cuba** at the end of the 1800s. Hearst himself went to Cuba. Hearst's papers had a lot to do with setting the stage for America going to war with Spain in Cuba.

As shown, the exact words written on the board appeared in boldface print throughout the scripts to prompt the instructor to write each of the 40 targeted facts on the board. Three additional facts were also written on the board to mirror typical teacher presentations in which all material presented in lecture is not included on the test. Finally, both scripts contained five prompts to encourage student involvement. These prompts consisted of questions asking students to give examples of current newspapers and tabloids, define basic words, or provide examples of languages other than English in which newspapers might be written today.

The scripts differed during the review portion of the lesson for the two groups of students. For the control group, statements focused on repetition of the reviewed facts. Specifically, for each reviewed fact, the instructor (a) verbally cued the importance of the fact, (b) prompted the students to take notes on the fact, (c) asked a question to elicit information about the fact, (d) covered the information again if the students did not remember the information, (e) wrote the information on the board, and (f) discussed the information with the students emphasizing key linkages/relations. (See Table 4 for an example script used for this review sequence.)

For the experimental group, the review portion included statements corresponding to the steps of the Recall Enhancement Routine for each reviewed fact—that is, cueing the importance of the fact, cueing notetaking, naming the mnemonic device, presenting the mnemonic device, and reviewing the device. The scripts for the two reviews were designed to include approximately the same number of statements and to take the same amount of time to present (i.e., 10 min). Both scripts covered the review of the same set of 21 reviewed facts. Thus, only the content of the review portion of the lesson was different in the two scripts.

Measurement Systems

Recall test. A multiple-choice test was constructed to measure student recall of the 40 targeted facts. A multiple-choice item comprised of a question and four optional

[1]One of the judges held a BA and had over 60 hr of graduate study and extensive experience in teaching and curriculum development. One held a PhD in special education and has been certified to teach at the secondary level. Another held BAs in education and special education, MAs in special education and counseling, and was completing work on a PhD in special education. The fourth judge had expertise in test construction and assessment and was pursuing a graduate degree in educational psychology and research.

TABLE 3
Use of Devices for Remembering

Information	*Appropriate Device*	*Example*
"**Tabloids** are characterized by the following: They contain many **pictures**, they are meant to be **interesting** and **gossipy**, and they are small in **size**."	Acronym device (first-letter mnemonic)	The first letter from each of the four words in the list (i.e., *pictures*, *interesting*, *gossipy*, and *small in size*) can be isolated, and an acronym can be formed. This results in the letters *p*, *i*, *g*, and *s*, which spells *PIGS*. To help remember the connection, the sentence, "Any day now, I expect to see *PIGS* in *TABLOIDS*" can be constructed.
"The **Copperheads** supported the **Confederacy**."	Mental imagery device (picture)	A mnemonic device can be constructed by forming a very strong mental image that links the Copperheads with the Confederacy. For example, the students might form a mental image of a statue of a soldier waving a Confederate flag. The head of the statue is copper. The copper shines brightly as the Confederate flag waves. The students construct a strong relation between the copperhead and the sign of the Confederacy.
"**Hearst** covered the war in **Cuba**."	"Boxing" device (keyword)	From the name *Hearst*, isolate, by drawing a box around the letters, *hears*, which sounds like a *hearse*, a large black car seen in funeral processions. Then, isolate the first syllable from *Cuba*, which is *cube*. Combine these two into a mental image of a large, black shiny hearse with a large cube of ice sitting on top. The word *Hearst* cues the students to visualize the hearse; the image of the hearse is combined with the *cube* to cue the linked word *Cuba*.

TABLE 4
Example Script for Repetition of Reviewed Facts

Importance cued	"You'll need to remember the elements of the code of ethics for newspaper editors."
Notetaking cued	"Be sure to get each pair or set of facts in your notes."
Information elicited	"Who can tell us the elements included in the code of ethics for newspaper editors?" (Teacher writes "**Code of Ethics: Fair, Accurate, Impartial, Responsible**" on the board.) Information was presented by teacher if students could not answer.
Information discussed	"Let's discuss these four important elements in the newspaper editors' code of ethics." (Teacher discusses with the students why editors need to be fair, accurate, impartial, and responsible.)
Information reviewed	"What will you remember about the newspaper editors' code of ethics?" (Elicited from the students: "The newspaper editors' code of ethics encourages editors to be fair, accurate, impartial, and responsible.")

responses was constructed to measure recall of each targeted fact. For example, the following question measured a student's recall of one of the elements in the code of ethics for newspaper editors:

> Which of the following was a guideline that newspaper editors put in their Code of Ethics in 1923 to curb abuses in their profession?
>
> (a) inexpensive
> (b) illuminating
> (c) impartial
> (d) idealistic

The test was administered to a group of 20 junior high students enrolled at a different school than the participants in our study. (These students received the lecture portion of the lesson before taking the test; they did not receive the review portion.) After these 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 each item. The index was used to assign facts to the reviewed or nonreviewed sets of facts as just described. Internal consistency for the instrument was calculated using Cronbach's coefficient alpha. An alpha of .87 was found for the sample.

In addition, content validity for the instrument was established using the same panel of four judges who were asked to read the lecture script and match the information in the lecture to the items on the test. All panel members determined that every item on the test was covered in the lecture. Once its validity and reliability had been established, the test was used to yield two measures: the percentage of correct responses on items corresponding to reviewed facts and the percentage of correct responses on items corresponding to nonreviewed facts.

Student notes. Notes taken by the students during the lecture were scored to determine whether they included critical information presented in the lecture. Two scores were calculated. To determine a quantity score for each student, the student's notes were compared to a check list of bits of information related to the 43 facts (40 test facts and 3 distractors) that were written on the blackboard by the teacher during the lesson. One point was awarded for each bit of information represented in the student's notes. For example, if the lecturer wrote "Thomas Fortune—early Afro-American editor" on the board, and the student's notes contained "Fortune" or "Thomas Fortune," the student was awarded 1 quantity point; if the student noted both "Fortune" and "Afro-American editor," 2 quantity points were awarded. A total of 70 pieces of information were possible. A student's quantity score was the percentage of individual pieces of information noted.

A quality score was also determined for each set of notes, again based on a comparison of the student's notes and the list of information written on the board. For this score, the student received 1 point for each set of items that was completely noted. A set of items consisted of at least two pieces of information that were presented together; however, in the case of lists of items, sets consisted of as many as five bits

of information. For example, if the student in the example just presented had noted both "Fortune" and "Afro-American editor" in his or her notes and had linked the two bits of information in some way (e.g., with a hyphen), that student was awarded 1 point for that set of information. If, however, either or both of the bits of information or the linkage between them was omitted, then the student received a score of 0. In other words, for the student to receive 1 point for the quality score, all of the bits of information to be associated together regardless of the number of items in that set had to be recorded in the student's notes in a way such that the information was connected. There was a total of 29 information sets written on the board.[2] A student's quality score was the percentage of information sets noted.

In addition, we also incorporated a measure of the information included in student notes that had been presented by the teacher during the lesson but had not been cued or written on the board. Thus, when an additional item appeared in a student's notes, the observer checked the script to determine whether it appeared in the lecture and, if it did, wrote the item verbatim on the score sheet. Raw scores of extra items were tallied for each student.

Interscorer reliability. Interscorer reliability was determined by having two scorers independently score a random sample (15%) of the tests and student notes. The two observers' 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. For the tests, the scorers agreed 199 times out of 200 opportunities (percentage of agreement = 99.5%). For the notes, the scorers agreed on 334 out of 348 opportunities (percentage of agreement = 96%). On individual sets of notes, the percentage of agreement ranged from 91% to 100%.

Procedures

The lesson was presented to the two groups in regularly scheduled class sessions. The instructor for both groups was the first author, who holds a PhD in special education and has been certified as an English teacher at the secondary level. The appropriate script was used for each group; although every effort was made to deliver the material naturally, the script was followed. To assure procedural integrity, a second researcher was present in the classroom during both sessions. The sessions were audiotaped, and the second researcher took notes, including a replication of all information written on the chalkboard, and ascertained that all information presented to the experimental group was also presented to the control group.

At the end of the lesson, student notes were collected. Both groups spent the same length of time on the lecture portion (45 min) and on the review portion of the lesson (10 min). On the next day, all students were administered the test. Test instructions were read to the students; however, test items were not read to the students. Students were given 45 min to complete the test. As they took the test, they were not allowed to consult their notes or each other.

[2]These information sets included lists of as many as five facts; thus, the discrepancy between 43 pieces of information and 29 information sets is explained.

Experimental Design

The experimental design is shown in Table 5. Several comparisons were identified as critical to answering the question whether implementation of the Recall Enhancement Routine had positive effects on student performance. For the first comparison, the experimental group students' test scores on reviewed facts were compared to those of the control group students to determine if the two types of reviews had differential effects. In a second comparison, the experimental group students' test scores on nonreviewed facts were compared to those of the control group students to validate the similarity of the two groups. Even though students had been randomly assigned to the groups, this comparison was incorporated because of the sample size. Thus, two types of experimental control were inherent in the design: random assignment and substantiation of equivalence of the groups by comparing performance on nonreviewed facts. Other comparisons focused on the percentage of students who performed at a level generally deemed as passing on classroom tests, student recall of repetitiously reviewed facts versus nonreviewed facts, differences between the performance of LD and NLD students, and student note-taking performance.

RESULTS

Test Results

Nonreviewed facts comparisons. As indicated, even though the students had been randomly assigned to the experimental or the control group, a comparison was carried out to substantiate that the two groups were indeed similar. Therefore, the students' test scores on nonreviewed facts were compared to verify that the groups performed similarly on items that were not reviewed in any way. Figure 1 shows the results of this comparison, indicating that the control LD students earned a mean score of 45.61% (SD = 15.57) correct, whereas the experimental LD students earned a mean score of 47.95% (SD = 14.99) correct on nonreviewed facts. In comparison, the control NLD students and the experimental NLD students earned mean scores of 64.47% (SD = 16.81) and 64.11% (SD = 15.93), respectively. The group mean score on nonreviewed facts for the LD and NLD students in the experimental group combined was 56.84% (SD = 17.21); for the combined LD and NLD students in the control group, it was 56.39% (SD = 18.54).

Further analysis was conducted to determine levels of student performance on nonreviewed facts when judged by

TABLE 5
Design

	Experimental Group		*Control Group*	
	LD[a]	*NLD*[b]	*LD*[a]	*NLD*[c]
Nonreviewed facts	Presented in lecture	Presented in lecture	Presented in lecture	Presented in lecture
Reviewed facts	Enhanced with routine	Enhanced with routine	Repeated	Repeated

[a]n = 9. [b]n = 11. [c]n = 12.

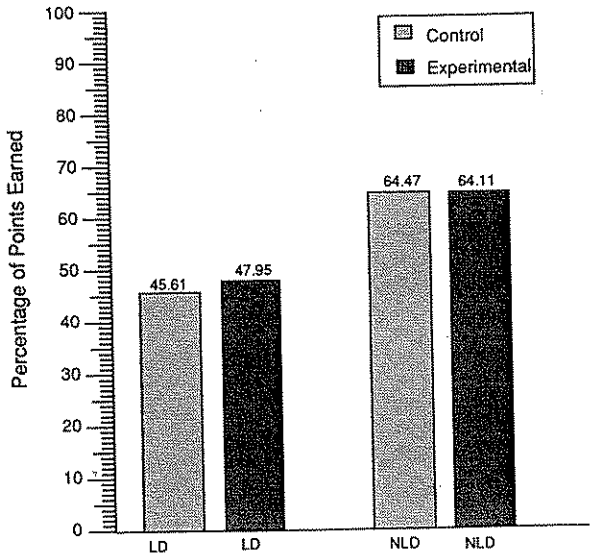

FIGURE 1 Student performance on nonreviewed facts.

standards that are often applied to content test performance in regular secondary classes. That is, results were analyzed to determine the percentage of students who would have performed at a level generally deemed as passing (i.e., a score of 60% or above) on classroom tests. The left side of Figure 2 indicates that the students' scores represent passing grades (i.e., scores above 60%) for 22% of the LD students in both control and experimental groups and for 58% of the NLD students in the control group and 63% of the NLD students in the experimental group. Thus, on the nonreviewed facts, the majority of the LD students and a substantial proportion of the NLD students scored within the failing range according to typical grading standards.

A univariate analysis of variance (ANOVA) was conducted on the nonreviewed facts data. Results are shown in Table 6. The only significant difference was related to the exceptionality dimension ($p = .001$): NLD students earned significantly higher scores on nonreviewed facts than the LD students. No difference was found between the experimental and the control groups. Thus, the two groups were considered to be comparable because they responded in similar ways to questions about nonreviewed facts on the test.

TABLE 6
Univariate Test Results for Nonreviewed Facts

Effect	*F Value*	*Significance*
Group × Exceptionality	.072	.789
Exceptionality	12.197	.001
Group	.039	.845

Reviewed facts comparisons. Figure 3 shows the test results on reviewed facts—that is, the effects of the Recall Enhancement Routine versus repetitious review on students' test scores. For the LD students in the control group, the mean test score was 41.80% ($SD = 20.17$), compared to 70.90% ($SD = 17.89$) for the LD students in the experimental group. For the NLD students in the control group, the mean test score was 64.29% ($SD = 20.86$), compared to 84.85% ($SD = 12.57$) for the NLD students in the experimental group. The mean test score for the LD and NLD students combined in the experimental group was 78.57% ($SD = 16.39$). For the whole control group, LD and NLD combined, the mean test score was 54.65% ($SD = 23.07$).

Another analysis examined the degree to which the intervention affected the percentage of students who received a passing grade on the test (i.e., a score of 60% or more). The results of this analysis are shown in Figure 2. The right side shows that 11% of the LD students in the control group earned passing scores on reviewed facts. For comparison, 77% of the LD students in the experimental group earned passing scores. Figure 2 also shows that 66% of the NLD students in the control group earned passing scores and that 100% of the NLD students in the experimental group earned passing scores on reviewed facts. Thus, all of the NLD students and a majority of the LD students earned passing scores when the information was enhanced through use of the routine and mnemonic devices.

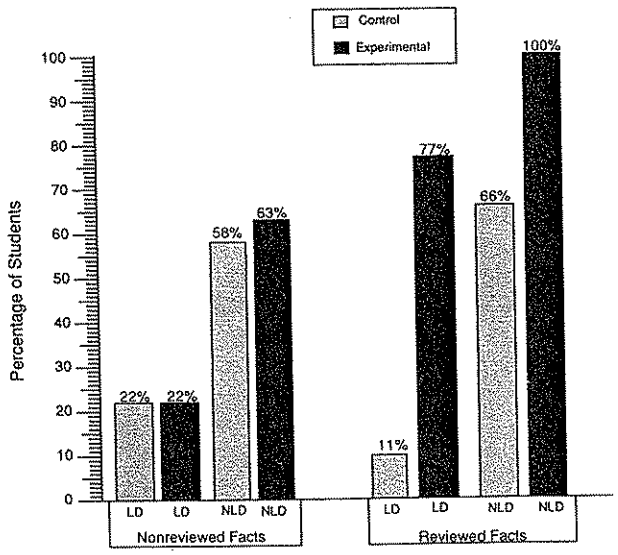

FIGURE 2 Percentage of students performing at passing levels.

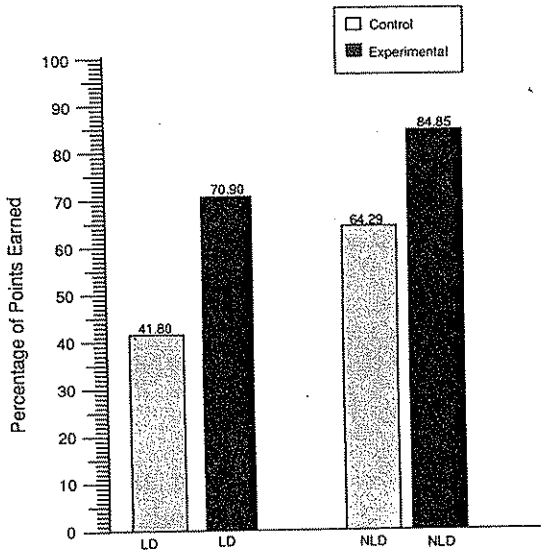

FIGURE 3 Student performance on reviewed facts.

TABLE 7
Univariate Test Results for Reviewed Facts

Effect	*F Value*	*Significance*
Group × Exceptionality	.559	.460
Exceptionality	10.173	.003
Group	18.900	.000

The results of the univariate ANOVA on the reviewed facts data are shown in Table 7. Significant differences were found for both exceptionality ($p = .003$) and group dimensions ($p = .000$). Specifically, NLD students scored significantly higher than LD students, and experimental group students scored significantly higher than control group students on items related to reviewed facts.

Nonreviewed versus repeated facts comparisons. A comparison of the data in Figures 1 and 3 reveals how the control group students performed on items related to repetitiously reviewed facts versus nonreviewed facts, thereby indicating the impact of repetitious review on students' test scores. As shown, LD students' mean scores on repetitiously reviewed facts were slightly less than their mean score on nonreviewed facts (41.80% vs. 45.61%). Furthermore, the difference in the NLD students' average scores on repetitiously reviewed facts and nonreviewed facts was minimal (64.29% vs. 64.47%). Overall, as shown in Figure 2, only 11% of the LD students' scores and 66% of the NLD students' scores on repetitiously reviewed facts were in the passing range. Thus, the instructional practice of using repetitious review to promote recall is not supported.

Note-Taking Results

The note-taking results are shown in Figure 4.[3] For the information presented during the lecture portion of the lesson, the LD students in the control group earned a mean quality score of 90%, whereas the LD students in the experimental group earned a mean quality score of 87% on their notes. Of the LD students in the control group, 62% took extra notes, compared to 66% of the LD students in the experimental group. The NLD students in the control group earned a mean quality score of 84%, whereas the NLD students in the experimental group earned a mean quality score of 96%. (The score of 84% includes the score of a student who normally did not take notes and who did not take notes during the lesson; when the score was recomputed without this student, the mean quality score for the NLD students in the control group was 93%.) Finally, of the NLD students in the control group, 80% took extra notes, compared to 70% in the experimental group. For the students who took notes on items that had not been written on the board, the NLD students in both the experimental and control groups took approximately twice as many notes as the LD students.

The LD students in the control group earned a mean quantity score of 89%, and the LD students in the experimental group earned a mean quantity score of 90%. By comparison, the NLD students in the control group earned a mean quantity score of 84% (93% without the student who did not take notes), and the NLD students in the experimental group earned a mean quantity score of 96%.

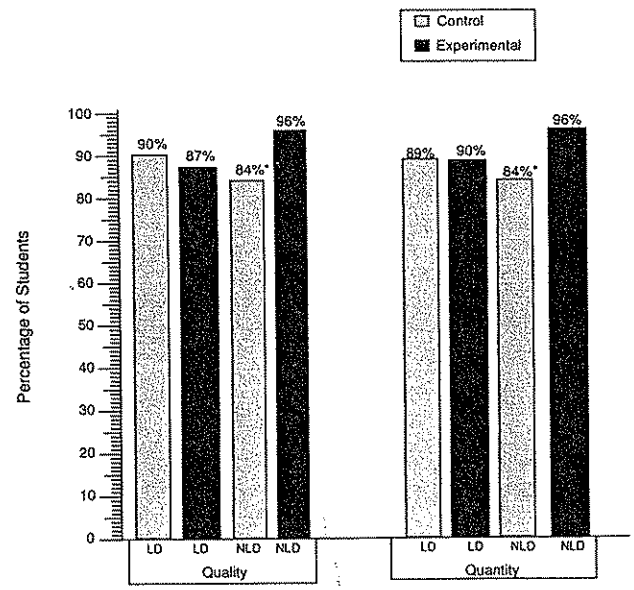

FIGURE 4 Student performance on note taking. *Average includes one student who did not take notes.

The students took few notes during the review portion of the lesson. Specifically, during the review, the LD students in the experimental group earned a mean quantity score of 3.82% and a mean quality score of 2.56% on their notes. Of the LD students in the experimental group, only two took notes during the review on the memory devices. In comparison, the NLD students in the experimental group earned a mean quantity score of 15.94% and a mean quality score of 16.15% during the review. Of the NLD students in the experimental group, only one student took notes on the memory devices; however, another student circled reviewed items and yet another student starred reviewed items. None of the LD and NLD students in the control group took notes during the review portion of the lesson.

DISCUSSION

In summary, this study resulted in several conclusions regarding the use of the Recall Enhancement Routine. First, the use of the Recall Enhancement Routine to facilitate recall of factual information for both LD and NLD adolescents in secondary regular classrooms yielded higher test scores than a traditional repetitious review of facts. The recall performance of both the LD and NLD students in the experimental group was substantially higher (by 29.10 and 20.56 percentage points, respectively) than the performance of similar students in the control group on reviewed facts. The substantial differences evidenced by both LD and NLD students in the experimental group are important because teachers have reported that if the magnitude of change evidenced by various subgroups (e.g., the high achievers, the low achievers, etc.) within a regular class is not clear and valued by students from each subgroup, they will not be supportive

[3]Notes from 38 of the 41 students were available. Therefore, note-taking data are reported for 8 LD and 10 NLD students in the control groups and 9 LD and 10 NLD students in the experimental groups. One NLD student in the control group turned in a blank sheet of paper, indicating that he did not normally take notes.

of the alternative teaching procedure used by the teacher. Once the teacher detects this dissatisfaction, he or she tends to discontinue the new instructional procedure (Schumaker et al., 1991). Because this instructional procedure described here can affect the performance of both LD and NLD students, it has potential for use by regular classroom teachers to facilitate instruction of academically diverse classes of students.

Second, application of the Recall Enhancement Routine markedly affected the percentage of students who scored in the passing range. That is, 100% of the NLD students and 77% of the LD students in the experimental group received passing grades on a classroom-type test over the content information, whereas a substantially smaller number of the students in the control group earned passing grades. This suggests that inclusion of mnemonic devices into content lessons may directly impact successful classroom test performance. This is critical because classroom teachers are not likely to continue using an educational innovation if substantial gains are not readily apparent on measures that they deem to be important (Fullan, 1991). The magnitude of change in test performance supports adoption and sustained use of the procedure.

Third, the LD students in the experimental group took only slightly better notes than the LD students in the control group (a difference between 1% and 3%); in addition, all the LD students had most of the facts represented in their notes. This finding replicates previous work (Bulgren et al., 1988) indicating that LD students appear to include in their notes what the teacher writes on the chalkboard. Similarly, the NLD students had most of the facts represented in their notes. Few students took notes about the mnemonic devices, however. These findings indicate that (a) the note-taking performance of LD students may not be markedly dissimilar from that of NLD students as long as teachers carefully write important information on the board, (b) the test performance differences found in this study for LD students are not related to differences in note-taking performance, and (c) LD students do not need to write the mnemonic devices in their notes in order to benefit from them in the short term. Because the test was administered in this study on the day after the lesson, further research is warranted to determine the relation between taking notes on mnemonic devices and student performance when the test is given several days or even weeks after the information is presented.

This study adds credence to the existing literature base on content enhancement routines (e.g., Bos & Anders, 1990; Bulgren et al., 1988; Lenz et al., 1987; Lovitt, Rudsit, Jenkins, Pious, & Benedetti, 1985; Schumaker et al., 1991) showing that regular classroom teachers can alter their instructional practices with their entire class in a way that improves the performance of both LD and NLD students. Thus, these data support the notion that including students with LDs in mainstream content classes at the secondary level can be successful if the delivery of curriculum content is enhanced.

Several limitations apply to this study. Most important, we did not address the practicality of incorporating the Recall Enhancement Routine into secondary lessons on an ongoing basis by including such questions as: How much planning and extra preparation is required? How satisfactory would teachers perceive the routine to be in facilitating their efforts to teach large amounts of content to a predetermined level of mastery? How readily can classroom teachers create remembering devices and, from an instructional design standpoint, how effective are such teacher-designed devices?

In addition, we did not address whether or not students can be taught to generate their own mnemonic devices within mainstream settings. If an overriding goal of inclusionary placements for students with LDs is to increase their ability to function as independent learners and performers in mainstream settings, addressing this issue is important. The data from this study are encouraging; students with LDs can be successful in a regular class setting. Nonetheless, they represent the effects of instructional accommodations. The measure of true integration of students with LDs in mainstream environments is whether these students can meet criterion levels of performance (e.g., on classroom tests, on measures of social acceptance, on measures of group participation, etc.) when the classroom teacher does not make accommodations for the student. Therefore, the pressing matter to be addressed is whether over time (e.g., several months of a school year) students with LDs can learn to generate mnemonic devices similar to those that have been used and modeled repeatedly by the classroom teacher and use those devices to respond successfully to classroom demands as independent learners and performers.

A final concern and an area for future research relate to the levels at which LD students can perform in the mainstream secondary classroom. Although this study and others have shown that more LD students can score in the passing range if content enhancement routines are used, a proportion of LD students (about one quarter) remain in the failing range, whereas others are barely passing. These findings are cause for concern and should act as springboards for future research focusing on questions such as: What constitutes "real" success in regular classes and What kinds of interventions or combinations of interventions are instrumental in producing "true" success? Only when students with LDs are performing successfully will they continue their educational experience and thereby become productive members of society.

ACKNOWLEDGMENTS

Preparation of this article was supported in part by the U.S. Office of Education (Office of Special Education Programs) Grant H133G90156.

We acknowledge the administrators and teachers of South Junior High School in Lawrence #497 School District, including Robert L. Taylor, Paulette Strong, P. Kay Duncan, Vicki Beals, Mary Ann Henry, Diane Low, and Denise Akins.

REFERENCES

Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York: David McKay.

Bos, C. S., & Anders, P. L. (1990). Effects of interactive vocabulary instruction on the vocabulary learning and reading comprehension of junior-high learning disabled students. *Learning Disability Quarterly, 13*, 31–42.

Bulgren, J., Deshler, D. D., & Schumaker, J. B. (1993). *Teacher use of a recall enhancement routine in secondary content classrooms*. Unpublished manuscript, University of Kansas, Center for Research on Learning, Lawrence.

Bulgren, J., Hock, M., Schumaker, J. B., & Deshler, D. D. (1993). Effects of teaching a strategy for remembering to students with learning disabilities. Manuscript submitted for publication.

Bulgren, J., Schumaker, J. B., & Deshler, D. D. (1988). The effectiveness of a concept teaching routine in enhancing the performance of students with learning disabilities in secondary mainstream classes. *Learning Disability Quarterly, 11*, 3–17.

Condus, M. M., Marshall, K. J., & Miller, S. R. (1986). Effects of the keyword mnemonic strategy on vocabulary acquisition and maintenance by learning disabled children. *Journal of Learning Disabilities, 19*, 606–613.

Deshler, D. D., & Schumaker, J. B. (1988). An instructional model for teaching students how to learn. In J. E. Zins & M. J. Curtis (Eds.), *Alternative educational delivery systems: Enhancing instructional options for all students* (pp. 391–411). Washington, DC: National Association of School Psychologists.

Edgar, E. (1993, March). *Curricula options at the secondary level: Preparing youth for the twenty-first century.* Paper presented at the Blazing New Trails Specific Learning Disabilities State of Minnesota conference, Brainerd, MN.

Ferro, S. C., & Pressley, M. G. (1991). Imagery generation by learning disabled and average-achieving 11- to 13-year-olds. *Learning Disability Quarterly, 14*, 231–239.

Fullan, M. (1991). *The new meaning of educational change.* New York: Teachers College Press.

King-Sears, M. E., Mercer, C. D., & Sindelar, P. T. (1992). Toward independence with keyword mnemonics: A strategy for science vocabulary instruction. *Remedial and Special Education, 13*, 22–33.

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. (in press). Teaching in the content areas. In P. A. Cegelka & W. H. Berdine (Eds.), *Effective instruction for students with learning problems.* Needham Heights, MA: Allyn & Bacon.

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

Lovitt, T., Rudsit, J., Jenkins, J., Pious, C., & Benedetti, D. (1985). Two methods of adapting science materials for learning disabled and regular seventh graders. *Learning Disability Quarterly, 8*, 275–285.

Mastropieri, M. A., & Scruggs, T. E. (1988). Increasing LD students' content area learning: Research implementation. *Learning Disability Research, 4*, 17–25.

Mastropieri, M. A., Scruggs, T. E., McLoone, B., & Levin, J. R. (1985). Facilitating learning disabled students' acquisition of science classifications. *Learning Disability Quarterly, 8*, 299–309.

Poplin, M. (1992). Educating in diversity. *American School Board Journal, 179*(3), 18–24.

Putnam, L. M., Deshler, D. D., & Schumaker, J. B. (1992). The investigation of setting demands: A missing link in learning strategies 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.

Schumaker, J. B., & Deshler, D. D. (1988). Implementing the regular education initiative in secondary schools: A different ballgame. *Journal of Learning Disabilities, 21*, 36–42.

Schumaker, J. B., Deshler, D. D., & McKnight, P. C. (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.

Scruggs, T. E., & Mastropieri, M. A. (1989). Mnemonic instruction of LD students: A field-based evaluation. *Learning Disability Quarterly, 13*, 119–125.

Scruggs, T. E., & Mastropieri, M. A. (1990a). The case for mnemonic instruction: From laboratory research to classroom applications. *The Journal of Special Education, 24*, 7–32.

Scruggs, T. E., & Mastropieri, M. A. (1990b). Mnemonic instruction for students with learning disabilities: What it is and what it does. *Learning Disability Quarterly, 13*, 271–280.

U.S. Department of Education. (1991). *America 2000: An education strategy sourcebook.* Washington, DC: U.S. Government Printing Office.