JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2006, **39,** 49–61 NUMBER 1 (SPRING 2006)

A COMPARISON OF INTERTEACHING AND LECTURE IN THE COLLEGE CLASSROOM

BRYAN K. SAVILLE AND TRACY E. ZINN

JAMES MADISON UNIVERSITY

AND

NANCY A. NEEF, RENEE VAN NORMAN, AND SUMMER J. FERRERI

THE OHIO STATE UNIVERSITY

Interteaching is a new method of classroom instruction that is based on behavioral principles but offers more flexibility than other behaviorally based methods. We examined the effectiveness of interteaching relative to a traditional form of classroom instruction—the lecture. In Study 1, participants in a graduate course in special education took short quizzes after alternating conditions of interteaching and lecture. Quiz scores following interteaching were higher than quiz scores following lecture, although both methods improved performance relative to pretest measures. In Study 2, we also alternated interteaching and lecture but counterbalanced the conditions across two sections of an undergraduate research methods class. After each unit of information, participants from both sections took the same test. Again, test scores following interteaching were higher than test scores following lecture. In addition, students correctly answered more interteaching-based questions than lecture-based questions on a cumulative final test. In both studies, the majority of students reported a preference for interteaching relative to traditional lecture. In sum, the results suggest that interteaching may be an effective alternative to traditional lecture-based methods of instruction.

DESCRIPTORS: interteaching, college instruction, education

Behavioral approaches to classroom instruction—including precision teaching (e.g., Lindsley, 1964), programmed learning (e.g., Skinner, 1968), direct instruction (e.g., Engelmann & Carnine, 1982), and Keller's (1968) personalized system of instruction (PSI)—have been available since the 1950s when Skinner (1954/1999) first discussed the application of behavioral principles in the classroom. Over the next 20 years, these approaches to classroom instruction were the focus of considerable experimental scrutiny. By the late 1970s, the literature was replete with publications reporting the success of behavioral approaches to classroom instruction (see Binder & Watkins, 1990; Buskist, Cush, & DeGrandpre, 1991; Moran & Malott, 2004). In the last 25 years, however, there has been a general decline in the use of these methods (e.g., Buskist et al.; Fox, 2004; Lamal, 1984; M. E. Lloyd & Lloyd, 1986). Possible reasons for this decline are that (a) some educators may be hesitant to adopt instructional methods that run counter to traditional pedagogical systems; (b) some educators have lamented that behavioral methods such as PSI do not fit nicely into the traditional academic calendar; (c) these methods typically require extensive preparation and resources, which may discourage some instructors from adopting them; and (d) misapplications or misunderstandings of behavioral principles have led some to conclude that these methods are ineffective. Given the reported effectiveness of behaviorally based methods of instruction, however, it is unfortunate that many educators are hesitant to implement these methods.

Recently, Boyce and Hineline (2002) introduced interteaching, a new method of

Address correspondence to Bryan K. Saville, Department of Psychology, MSC 7401, James Madison University, Harrisonburg, Virginia 22807 (e-mail: savillbk@jmu.edu).

doi: 10.1901/jaba.2006.42-05

classroom instruction that preserves certain elements of earlier behavioral approaches but offers more flexibility for teachers. Interteaching includes elements of PSI and precision teaching as well as elements from other empirically based instructional methods such as reciprocal peer tutoring (Griffin & Griffin, 1998), cooperative learning (e.g., Halpern, 2004), and problem-based learning (e.g., Duch, Groh, & Allen, 2001). Moreover, interteaching attempts to remedy several of the purported problems with other behavioral approaches to classroom instruction (Boyce & Hineline, 2002; Saville, Zinn, & Elliott, 2005). For example, because interteaching, unlike PSI, does not require students to master course content, grade distributions are less likely to be negatively skewed. Also, interteaching allows students to work at their own pace during each class session but incorporates deadlines for the completion of course assignments, which may obviate student procrastination. Finally, although initial course preparation may be time consuming, the focus on student discussion ultimately reduces the amount of time spent preparing lectures during the semester. Moreover, the resulting lectures tend to be brief and are focused primarily on points of difficulty or student interest.

The sequence of a typical class period using the interteaching method proceeds as follows. The instructor first constructs a preparation (prep) guide that contains questions designed to guide students through a specific unit of course material. Students then use the prep guide to prepare for the subsequent class period. During the class period, students form small groups, typically dyads or triads (and usually composed of different students each session), and discuss the questions, for which they receive a small amount of course credit. Concomitantly, the instructor (and teaching assistant, if necessary and available) moves among groups, answering questions, monitoring students' comprehension of the material, and in general facilitating group discussion. At the end of each session, students fill out an interteaching record, the purpose of which is to help the instructor identify questions that posed difficulties for students. The instructor then uses this information to construct a short lecture designed to address any questions the students requested. The lecture begins the subsequent class period and precedes group discussion for that day.

There are additional components of the interteaching method. First, Boyce and Hineline (2002) recommended that there should be a clear connection between prep guide questions and test items. Specifically, they suggested that tests, or probes, should contain several essay questions that come directly from the prep guides; other test questions (e.g., fill in the blank, multiple choice, short answer) should closely mirror the types of questions found in the prep guides and covered in class. In addition, Boyce and Hineline suggested that instructors should give several probes (i.e., at least five) during the course of a semester.

In addition, Boyce and Hineline (2002) discussed two methods for improving the quality of group discussions: (a) Instructors can ask students to rate the quality of their discussions and provide an explanation for their ratings (e.g., ''My partner was not prepared to discuss the material''), and (b) instructors can introduce a performance contingency in which students receive ''quality points'' if both they and their partners do well on specific portions of the tests (e.g., if both receive high scores on a particular essay question). Finally, although not discussed by Boyce and Hineline, instructors could award additional quality points during class to students who engage in effective discussion. Regardless of the specific method, the reason for awarding quality points is to introduce a contingency whereby students quickly learn that effective discussion results in better course grades.

Although anecdotal evidence (Boyce & Hineline, 2002) supports the efficacy of interteaching, few empirical data exist. In a recent

study, Saville et al. (2005) examined the effectiveness of interteaching relative to other more traditional methods of instruction in a controlled laboratory setting. They randomly assigned 84 college students to one of four conditions—interteaching, lecture, reading, or a control condition. Participants in the interteaching group read a short journal article and participated in the activities described earlier. In the lecture and reading conditions, respectively, participants either heard a lecture over the same article or simply spent time reading the article. One week later, participants in each of the experimental conditions, as well as participants in the control condition who had no exposure to the article, took a 10-question multiple-choice quiz. The type of teaching method had a statistically significant effect on quiz scores: Participants in the interteaching condition correctly answered a greater percentage of quiz questions (75%) than did participants in the other groups. In addition, there was not a statistically significant difference among scores in the lecture (60%), reading (55%), and control (52%) conditions.

Although Saville et al.'s (2005) results provide empirical support for the efficacy of interteaching under highly controlled conditions, it is important to examine the generality of these effects in the more typical context of university courses. Thus, the purpose of the present studies was to examine the effectiveness of interteaching in relation to a more traditional form of instruction—lecture—in typical university classroom settings.

STUDY 1

METHOD

Participants and Setting

Participants in Study 1 were 35 graduate students, the majority of whom were in a 1-year MA degree program in special education; the remaining students were from graduate programs in psychology and physical education. We conducted the study across eight weekly class sessions of a course that focused on promoting transition from school to adult roles. Each session, which lasted 2 hr 15 min, began with a 30-min review and a 15-min quiz. Then, after a 15-min break, students either heard a lecture or engaged in group discussion (see below). Finally, students had approximately 10 min to submit questions based on the material covered in class. Any time remaining in each class session was allocated to supplemental activities not associated with the study (e.g., guest presenters who described their experiences or their agency's role in the transition process, discussion of class projects).

Materials and Procedure

To examine the effects of instructional method on students' weekly posttest performance (and to compare performance on posttests with pretests that preceded either instructional condition), we alternated lecture and interteaching across class sessions. Before the start of the course, we designated each of the eight class sessions as either lecture or interteaching, with the constraint that each condition must occur an equal number of times and for no more than two consecutive class sessions.

Lecture. During class sessions designated for the lecture condition, one of the two coinstructors for the course delivered a lecture with PowerPoint® slides. The slides were projected on a large screen and also were available as handouts (downloaded from the instructor's Web site). Each lecture lasted 40 to 60 min and began with a summary of the course objectives for the session, which were also specified on the course syllabus. The lecture then proceeded sequentially through each objective with corresponding material that reviewed and supplemented key points from the assigned readings. To encourage participation, students had 10 min at the end of each lecture to submit written questions, which counted as bonus points toward their course grades.

Interteaching. During sessions designated for interteaching, we divided students into dyads or

triads by asking them to work with another student in an adjacent seat; we alternated the grouping criterion (e.g., ''Work with the student to your right.'' ''Work with the student behind you.''), such that students typically worked with someone different each week. Students then had 40 to 60 min to discuss the questions contained on the prep guides. The guides were available before class on the instructor's Web site. They contained comprehension- and application-level questions (Bloom, 1956) based on the course objectives, and were designed to guide students through the reading assignments. While each group discussed the questions, two graduate teaching assistants, who were completing a college teaching practicum, circulated through the room and distributed tickets when students' discussions were focused on the topics contained in the prep guides. The tickets were exchangeable for a chance to win extra points in a drawing held at the end of each session. The assistants also helped students if they had questions or needed clarification on a specific topic. During this time, the instructor either provided written feedback on students' assignments or provided additional assistance to students if both assistants were occupied. At the conclusion of the group discussions, students had 10 min to complete an interteaching record distributed by the assistants. Students used this form to request further clarification on particular topics and to rate the quality of their discussions. In preparation for the next class, the assistants examined each interteaching record and used the information to construct a lecture designed to clarify and review any material with which students reported having difficulty.

Quizzes

The primary dependent measure was the mean number of correct responses across all students on the quizzes, each of which was worth six points. We allocated one point for each correct answer to a question (or part of a question when it called for more than one discrete response). All quiz questions, which were based on the course objectives, required short answers. For example, a question for the objective, ''Be able to define transition services (incorporating its essential elements) as it relates to the preparation of students with disabilities for successful adult roles'' consisted of the definition of transition services with blank spaces for students to insert the missing essential elements. A question for the objective, ''Be able to describe cultural differences that can affect transition processes and planning'' consisted of asking students to state how a specified cultural group traditionally viewed the cause of disabilities.

We administered eight pretests (one for each unit) at the beginning of the quarter to assess knowledge of the material before instruction and to demonstrate the degree of learning associated with the two instructional methods. We informed students that they would not be penalized for incorrect answers, but that points earned for correct answers to pretest questions would be applied to the corresponding posttest quiz (posttest points counted toward the students' course grades). The purpose of this contingency was to control for the effects of motivation on performance. We did not, however, inform students of whether their answers on the pretests were correct until after they had taken the corresponding posttest.

If a student missed a class period (which rarely occurred), we deleted his or her quiz score from the analysis. We recorded the posttest quiz scores before the addition of any points earned from the corresponding pretest.

Interobserver Agreement

The primary instructor and two graduate teaching assistants independently scored 31% of the quizzes from each of the experimental conditions (i.e., pretest, interteaching, lecture). We scored the answers from each quiz on a separate data sheet to ensure independent scoring. To determine interobserver agreement, we calculated occurrence plus nonoccurrence agreement scores by dividing the number of

agreements by the number of agreements plus disagreements and multiplying by 100%. Mean agreement scores were 96% for pretest (range, 73% to100%), 96% for interteaching (range, 92% to 98%), and 86% for lecture (range, 75% to 97%).

Social Validity

At the conclusion of the course, 33 students anonymously completed a questionnaire containing three questions. The first question asked students to rank their preferences for interteaching or lecture (strongly prefer interteaching, somewhat prefer interteaching, no preference, somewhat prefer lecture, or strongly prefer lecture). The second question asked students to rate the extent of learning with interteaching and lecture. The final question asked students to indicate the reasons for their preferences.

Results and Discussion

Quizzes. Figure 1 displays the mean scores on each of the eight pretests given at the beginning of the quarter, as well as the mean scores on the posttest quizzes that followed the four interteaching sessions (Quizzes 3, 4, 6, and 7) and the four lecture sessions (Quizzes 1, 2, 5, and 8). Mean posttest scores were consistently higher following sessions of interteaching ($M = 4.68$, $SD = 0.31$) than following sessions of lecture ($M = 3.32$, $SD = 1.11$). Moreover, although both methods of instruction produced improvements relative to pretest scores ($M = 0.34$, $SD = 0.3$), interteaching produced improvements that were consistently greater than the improvements produced by lecture.

Social validity. As the top panel of Table 1 shows, the majority of students reported that they preferred interteaching, with fewer students reporting that they preferred lecture or had no preference. In response to the second question, a plurality indicated that they learned equally well with the two instructional methods. However, nearly as many students indicated that they learned more with interteaching. A considerably smaller number of students reported that they learned better with lecture. In response to the third question (not shown in Table 1), students who reported a preference for interteaching often cited its usefulness as an instructional guide for test preparation (e.g., "Interteaching gives you a guide to the main points of chapters and helps direct the way you study"). In contrast, the smaller number of students who favored lecture typically indicated why they did not prefer interteaching rather than why they preferred lecture (e.g., "I feel that the discussions with interteaching were not adding any additional knowledge than from reading the text").

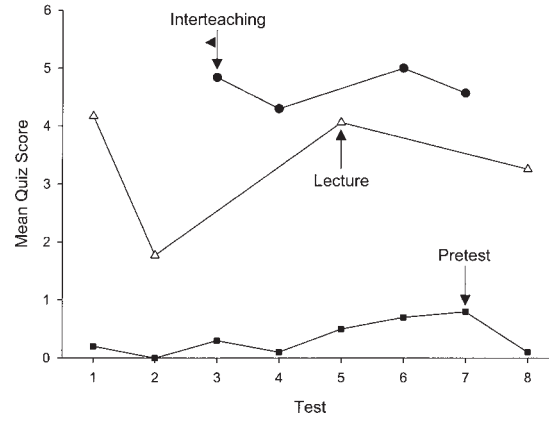

Figure 1. The mean number of questions answered correctly by students on each quiz in Study 1. Students took pretests at the beginning of the semester and posttests later in the semester after covering the corresponding course material. Filled squares represent pretests, filled circles represent quizzes taken after interteaching, and open triangles represent quizzes taken after lectures.

These results provide additional support for the notion that interteaching may be an effective alternative to lecture-based methods of instruction (Boyce & Hineline, 2002; Saville et al., 2005). In addition, our social validity data indicate that more students preferred the interteaching format. However, it is possible that higher quiz scores following interteaching sessions may have been a function of other uncontrolled variables such as the difficulty of course material covered on a given day. In

Table 1
Social Validity Data for Participants in Study 1 and Study 2

Question	Percentage (number) of students	
Study 1 (n = 33)		
1. Which did you prefer?		
Strongly prefer interteaching	0 (0)	
Somewhat prefer interteaching	58 (19)	
No preference	9 (3)	
Somewhat prefer lecture	18 (6)	
Strongly prefer lecture	15 (5)	
2. Extent of learning with interteaching and lecture		
Much more with interteaching	18 (6)	
Somewhat more with interteaching	18 (6)	
Equally well with interteaching and lecture	40 (13)	
Somewhat more with lecture	15 (5)	
Much more with lecture	9 (3)	
Study 2 (n = 27)		
	SEC 1	SEC 2
1. Which did you prefer?		
Interteaching	89 (8)	83 (15)
No preference	11 (1)	17 (3)
Lecture	0 (0)	0 (0)
2. With which did you learn more?		
Interteaching	78 (7)	78 (14)
No difference	22 (2)	17 (3)
Lecture	0 (0)	5 (1)

Study 2, therefore, we counterbalanced the two methods across two sections of an undergraduate research methods course to examine the generality of the findings.

STUDY 2

METHOD

Participants and Setting

Participants in Study 2 were 31 undergraduate students (5 men and 26 women). The participants' median age was 21 years (range, 19 to 32), and all were classified as either juniors (n = 23) or seniors (n = 8) in college. All participants were students enrolled in one of two sections of an undergraduate psychological research methods course. There were 12 participants (1 man and 11 women) in the first section (SEC 1) and 19 (4 men and 15 women) in the second section (SEC 2). SEC 1 met at 8:00 a.m. on Tuesdays and Thursdays for 75 min per class; SEC 2 met at 10:00 a.m. on Mondays, Wednesdays, and Fridays for 50 min per class.

Because we could not randomly assign participants to the different sections, we collected on the first day of class additional self-reported demographic data to determine the extent to which participants in the two sections were similar to one another. The obtained measures included (a) cumulative grade point average; (b) number of psychology courses previously taken; (c) number of credit hours taken during the semester; (d) grade in an undergraduate statistics course, which was a prerequisite for the research methods course; (e) whether students were employed and, if so, how many hours per week they worked; (f) whether students had significant others; and (g) whether they were members of fraternities or sororities. There were no statistically significant differences[1] between the two sections on any of the demographic measures.

[1] Although we conducted statistical analyses for all data in Study 2, we did not include the specific details of each analysis. All statistical analyses used an alpha level of .05, and any additional information regarding statistical analyses can be obtained from the first author.

Table 2
Order of Method of Instruction for Each Section in Study 2

Section	Method of instruction and number of classes preceding each test					
SEC 1	Lect (2)	Inter (3)	Lect (3)	Inter (4)	Lect (5)	Inter (3)
SEC 2	Inter (3)	Lect (5)	Inter (5)	Lect (5)	Inter (5)	Lect (5)

Note. Inter = interteaching; Lect = lecture.

Materials and Procedure

As in Study 1, we used an alternating treatments design (Kazdin, 1982), switching between interteaching and lecture several times throughout the semester. In addition, we counterbalanced the type of instruction across sections, such that while one section participated in interteaching, the other section heard a standard lecture over the same material (see Table 2).

Interteaching. The general procedure for interteaching was similar to that in Study 1, except for the following differences. First, for each discussion in which they participated, students received points that over the course of the semester amounted to approximately 8% of their final grades. In accord with Boyce and Hineline's (2002) suggestion, we required students to work with many different people during the semester; specifically, students were not to work with the same person more than three times during the semester. Second, the course instructors facilitated the discussions, and, although they did state at the beginning of each lecture class the general objectives for the day, they did not use PowerPoint® presentations, instead opting to write information on the blackboard as needed. Finally, rather than distributing quality points during the group discussions, we awarded quality points based on test performance (see below).

Lecture. During lectures, the instructors summarized the course objectives, introduced new material, and then solved various problems in an attempt to show students how to analyze and apply that material. To control for the content covered in class, the instructors lectured over the same information contained on the prep guides.

Tests

Unit tests. After each unit of information, students from both sections took identical tests that were graded by a graduate teaching assistant who was blind to the purpose of the study. Each test consisted of four essay questions taken from the prep guides (Boyce & Hineline, 2002) and worth approximately one third of the points. For example, one prep guide item asked students to "Discuss the five characteristics of small-*N* designs, and compare and contrast these characteristics with the characteristics of large-*N* designs." A related essay question on a subsequent test read,

> John and Sally are both interested in studying the effects of Vivarin® (caffeine) on memory. However, they differ with respect to the experimental approaches they will use to study the topic: John is going to use a small-*N* design, and Sally is going to use a large-*N* (i.e., between-groups) design. Describe five ways in which these studies will differ from one another.

The remainder of each test consisted of a variety of objective questions (e.g., fill in the blank, short answer), the majority of which required students to solve novel problems, apply information, or otherwise show higher level comprehension of the concepts contained on the prep guides or covered in lectures. For example, one prep guide question asked, "What are nine threats to internal validity? Explain how each of these would make it difficult to determine if the IV affected the DV," and a related test question required students to identify which threat to internal validity was

operating in the following scenario: "You are studying running speed in newborn rats to see if special rat food produces changes in motor efficiency. You measure their running speed on Day 2 and then again on Day 21." In total, students took six tests during the semester, three of which followed interteaching sessions and three of which followed lectures.

For each test, students in the interteaching condition also received quality points based on their essay answers. If, on a given essay question, both students who discussed that question during a given class period received either an A or a B (i.e., 4 or 5 points), each received four points toward his or her overall course grade. If, however, one or both students received less than a B (i.e., less than 4 points) on a given question, neither received additional quality points for that question. Thus, because each test contained four essay questions, students had the opportunity to earn 0 to 16 quality points toward their course grades. Overall, quality points accounted for approximately 8% of each student's final course grade.

Cumulative final test. A cumulative final test contained a combination of fill-in-the-blank, multiple-choice, and short-answer questions, similar to the types of objective questions contained on the unit tests. Of the 40 questions on the final, 20 were interteaching-based questions and 20 were lecture-based questions. Specifically, six questions came from information in the prep guides that preceded Test 1, six questions came from information in the prep guides that preceded Test 2, and seven questions each (i.e., 28 total) came from information in the prep guides that preceded Tests 3 through 6. However, it is important to remember that, because we alternated and counterbalanced conditions across classes, the lecture-based questions for SEC 1 were the same as the interteaching-based questions for SEC 2 and vice versa.

Social Validity

At the end of the semester, 27 students (SEC 1 = 9, SEC 2 = 18) responded anonymously to a three-item questionnaire designed to obtain their opinions regarding the effectiveness of interteaching and lecture. The first two questions asked the following: (a) Overall, which did you like more: interteaching or lecture? (b) Overall, did you feel you learned more with lecture or interteaching? For each of these questions, students had the option of circling interteaching, no preference, or lecture. Also, because we felt it important to determine if the instructors' different lecture styles affected the outcomes of the study, we asked students to rate on a scale from 1 (*very bad*) to 7 (*very good*) how good their instructor was at lecturing.

Results and Discussion

Unit test scores. Students' scores on each of the six unit tests was our primary dependent measure. As Figure 2 shows, test scores following interteaching sessions were typically higher than test scores following lectures, although they were not as pronounced as the differences observed in Study 1. On four of the six tests (1, 2, 4, and 6), students in the interteaching condition scored, on average, 12 percentage points higher than students who heard lectures over the same material. Moreover, although the difference was not as prominent on Tests 3 and 5, it tended in the same direction: On these tests, students in the interteaching condition scored approximately 3.5 percentage points higher than students in the lecture condition.

Although test scores following interteaching were typically higher than test scores following lecture, the difference between conditions was consistently larger when SEC 1 participated in interteaching and SEC 2 heard lectures (i.e., on Tests 2, 4, and 6). Conversely, the difference between conditions was smaller on two of three tests (Tests 3 and 5) when SEC 2 participated in interteaching and SEC 1 heard lectures. There are at least two possible reasons for this result. First, it is possible that the instructors' different teaching styles may have affected student

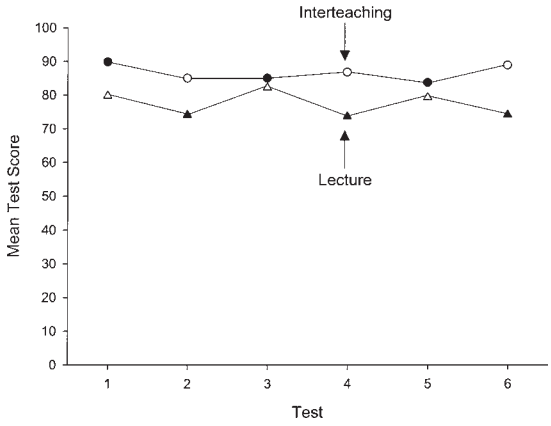

Figure 2. The mean percentage of questions answered correctly on each unit test in Study 2. Circles represent tests taken after interteaching, and triangles represent exams taken after lectures. Open symbols represent mean scores for SEC 1, and filled symbols represent mean scores for SEC 2.

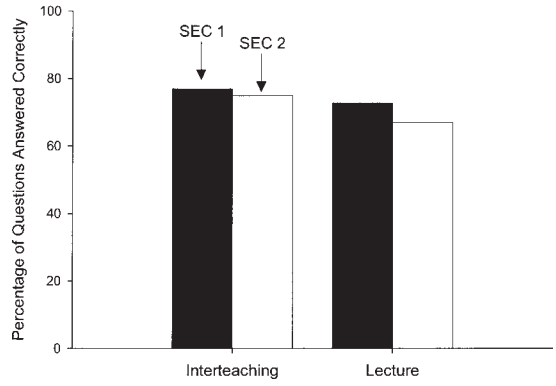

Figure 3. The mean percentage of interteaching-based and lecture-based questions answered correctly on the final exam. Filled bars represent the means for SEC 1, and open bars represent the means for SEC 2.

performance, resulting in the observed differences. Second, because course material also changed across tests, it is possible that interteaching and lecture might differentially affect learning depending on the type of material being covered.

Individual data (not shown) showed a pattern similar to the one observed in the group data for the majority of students. Specifically, 25 of the 31 participants (81%) typically scored higher following interteaching than following lectures. Of the remaining 6 participants, 5 scored approximately the same following interteaching and lectures; only 1 participant consistently did better following lectures.

Cumulative final test scores. As Figure 3 shows, students in both sections correctly answered a greater percentage of interteaching-based questions than lecture-based questions on the cumulative final test. Specifically, students in SEC 1 answered correctly 77% of the interteaching-based questions and 72% of the lecture-based questions; students in SEC 2 answered correctly 75% of the interteaching-based questions and 67% of the lecture-based questions. In addition, although this difference was more pronounced for students in SEC 2, the interaction between type of question and section was not statistically significant.

Social validity. As the lower portion of Table 1 shows, a large majority of students from both sections reported that they liked interteaching better than lecture. Conversely, no students reported that they preferred lecture, and only a small number reported that they had no preference. A large majority of students from both sections also reported that they felt they learned more with interteaching, whereas significantly smaller numbers reported that they learned the same amount with interteaching and lecture or learned more with lecture. Lastly, both instructors received median ratings of 7.0 on a 7-point scale (not shown), which suggested that students from both sections perceived their instructors as very good at lecturing.

In support of our results in Study 1 and in accordance with Saville et al.'s (2005) laboratory analysis, we observed that test scores following interteaching were typically higher than test scores following lectures. We also found that students correctly answered more interteaching-based questions on a cumulative final test. Moreover, students reported that they enjoyed interteaching more than lecture and learned more with interteaching. Finally, although students in both sections reported that

their instructors were very good at lecturing, test performance may have been affected by actual instructor performance in the classroom or changes in the difficulty of course material.

GENERAL DISCUSSION

The results of the present studies suggest that interteaching may be an effective alternative to the more traditional format of lecture-based instruction. There are several reasons why interteaching might be effective. First, interteaching involves group (typically dyadic) discussion, which seems to have numerous positive effects on learning. By participating in dyadic discussion, students engage in active learning, whereby they frequently practice the skills that form the behavioral repertoires teachers ultimately wish to shape. In numerous studies, researchers have found that active participation promotes learning (e.g., Mathie et al., 1993; Yoder & Hochevar, 2005).

In addition, the inclusion of dyadic discussion introduces several other reinforcement contingencies that may positively affect learning. For example, dyadic discussion capitalizes on immediate social consequences from peers (Boyce & Hineline, 2002), a variable that seems to enhance learning (Astin, 1993) but over which instructors have little control in traditional classroom settings (Michael, 1991). In addition, dyadic discussion actually seems to increase one-to-one interaction between students and instructor: Because the role of the instructor during discussion is to move among groups, facilitating discussion and answering questions, we have found that we actually interact with students more than we do during lectures. Whether called ''teacher approachability'' (e.g., Buskist, Sikorski, Buckley, & Saville, 2002; Schaefer, Epting, Zinn, & Buskist, 2003), ''student–teacher rapport'' (Buskist & Saville, 2004), or ''teacher immediacy'' (Wilson & Taylor, 2001), this interaction allows instructors to deliver additional reinforcers for appropriate behavior, thereby further affecting learning. Finally, dyadic discussion creates an implicit cooperative learning environment in which students work together to learn the course material, a practice that tends to increase performance relative to noncooperative contingencies (e.g., Johnson, Maruyama, Johnson, Nelson, & Skon, 1981).

Second, interteaching may be effective because offering prep guides to students helps create a clear link between study and test materials, thus increasing motivation, decreasing test anxiety, facilitating learning, and resulting in better retention of course material (e.g., Dickson, Miller, & Devoley, 2005; Flora & Logan, 1996). In fact, Michael (1991) stated that one of the biggest errors in college teaching is failing to create a clear link between course materials and test items. As such, Michael suggested providing students with weekly study objectives that help shape appropriate behavioral repertoires. With interteaching, providing students with prep guides likely sets the occasion for appropriate studying behavior and may help obviate the ''procrastination scallop'' (Michael, 1991) that often accompanies traditional lectures. Moreover, the similarity between prep guide questions and test items (Boyce & Hineline, 2002) is likely to exert stimulus control that may positively affect test-taking performance.

Third, whereas lectures typically function as antecedents, interteaching lectures follow studying behavior, thereby possibly serving a reinforcing function (Boyce & Hineline, 2002; K. E. Lloyd et al., 1972). Moreover, because students request information found in the interteaching lectures and because these lectures tend to be relatively short and focused on very specific material, the clarifying nature is likely to maintain the active interest of students, a feat that is sometimes hard to accomplish with traditional lectures (e.g., Benjamin, 2002; Boreham, 1984).

Finally, the inclusion of quality and participation points introduces additional reinforce-

ment contingencies that potentially affect learning. The inclusion of quality points, for example, creates an explicit cooperative contingency in which students' course grades are partially dependent on the behavior of other students, a condition that tends to improve performance. Similarly, the inclusion of participation points may function to increase class attendance (e.g., Beaulieu, 1984; Hansen, 1990), a factor that may positively affect learning (e.g., Shimoff & Catania, 2001).

Our experiences with interteaching also suggest that, for at least two reasons, instructors may enjoy interteaching more than traditional methods. First, although course preparation is initially time consuming, the focus on student discussion ultimately reduces the amount of time that instructors need to spend preparing additional course materials, an important consideration given the ever-increasing demands on faculty time (e.g., Milem, Berger, & Dey, 2000; Wright et al., 2004). Specifically, because instructors can quickly identify which prep guide questions are problematic for students, subsequent iterations of class material require significantly less work; and because students request the information contained in each clarifying lecture, instructors will not have to spend considerable time deciding what to include, a task that can be both frustrating and time consuming. Second, interteaching increases the amount of contact between instructor and students, a variable that may enhance student–teacher rapport (Buskist & Saville, 2004) and further enhance an instructor's positive attitudes toward teaching.

Although our results suggest that interteaching might be an effective alternative to lecture, there are several limitations to the present studies. A primary limitation is that we did not perform a component analysis to determine which elements of interteaching were responsible for its effectiveness. The purpose of the present studies was simply to determine if interteaching, as described by Boyce and Hineline (2002), was more effective than traditional lecture, arguably the most common method of instruction in college classrooms (e.g., Benjamin, 2002). However, because our manipulation of teaching method, in essence, included several different manipulations (e.g., access to prep guides, dyadic discussion, participation points), many of which have been shown in previous research to have positive effects on learning (e.g., Flora & Logan, 1996; Johnson et al., 1981; Shimoff & Catania, 2001), it is possible that the present studies contained one or more confounding variables and, consequently, low construct validity.

Another limitation was the difference in class length and, consequently, number of class periods for the sections in Study 2. Whereas SEC 1 met twice a week for 75 min each time, SEC 2 met for 50 min, three times a week. Because distributed practice tends to produce better performance than massed practice on a number of tasks (see Donovan & Radosevich, 1999), it is possible that the differences we observed were due to uncontrolled differences in class length. Although Boyce and Hineline (2002) suggested that interteaching might work better when classes meet more frequently, no systematic research has examined this possibility. Therefore, it would be interesting to determine exactly how class length affects the efficacy of interteaching.

A final limitation of the present studies concerns our interobserver agreement data. Although our agreement data for Study 1 are acceptable by conventional standards (e.g., Cooper, Heron, & Heward, 1987), mean agreement scores for pretests (96%) and for interteaching-based tests (96%) were considerably higher than the mean agreement score for lecture-based tests (86%), which may have contributed to differences in test scores. One potential reason for this discrepancy is that answers on the interteaching-based tests tended to be more precise, making it easier for the graders to identify answers that were similar to those on the grading key; similarly,

because few students answered questions on the pretests correctly, these items were also easy to grade. Conversely, students' answers on lecture-based tests were considerably more variable, which made grading more difficult and likely affected agreement scores. In Study 2, although the unit tests were graded by a teaching assistant who was blind to the purpose of the study, we failed to collect additional interobserver agreement data. Therefore, it is possible that some of the variability in test scores might be a function of unidentified changes in the way the assistant graded tests. Although grading the objective questions (i.e., fill-in-the-blank questions) was relatively straightforward, there were certain questions (e.g., essay questions) that were more subjective in nature and, therefore, subject to grading fluctuation.

In sum, the present results are encouraging. However, we believe that future researchers would do well to examine interteaching further. Specifically, we believe that it would be useful to identify which components of interteaching are necessary to produce positive changes in learning. For instance, to determine the extent to which access to prep guides contributed to differences in the present studies, one could replicate Study 2 but provide prep guides to students in the lecture condition. Similarly, one could examine the effects of quality points and participation points, respectively, by having two interteaching conditions but including quality points and participation points for only one of the conditions. Regardless, if future research finds that not all components of interteaching are necessary to produce changes in learning, then the construct of interteaching would necessarily be more circumscribed than Boyce and Hineline (2002) initially suggested.

In addition, we believe that researchers would do well to examine ways that interteaching could be more effective. For example, we believe that including an explicit contingency for completing prep guide questions might have positive effects on learning. Although anecdotal evidence suggests that most students answered the prep guide questions prior to class, it was not uncommon for students to report that their partners were not adequately prepared to discuss the material, which may have had a negative impact on learning. Furthermore, although there are several ways to improve the quality of group discussions, we believe that providing points contingent on effective discussion (as in Study 1) might have a more positive effect on learning than (a) having students report of the effectiveness of their discussions or (b) awarding points based on test scores (as in Study 2; see Boyce & Hineline, 2002).

We hope our results prompt other researchers interested in behavioral approaches to education to examine interteaching further. If future research validates interteaching as an effective alternative to traditional teaching methods, we hope that more college instructors will adopt it in their courses and introduce others to the efficacy, flexibility, and practicality of this new behavioral approach to classroom instruction.

REFERENCES

Astin, A. W. (1993). *What matters in college?: Four critical years revisited*. San Francisco: Jossey-Bass.

Beaulieu, R. P. (1984). The effects of traditional and alternate rewards on attendance. *College Student Journal*, *18*, 126–130.

Benjamin, L. T. Jr. (2002). Lecturing. In S. F. Davis & W. Buskist (Eds.), *The teaching of psychology: Essays in honor of Wilbert J. McKeachie and Charles L. Brewer* (pp. 57–67). Mahwah, NJ: Erlbaum.

Binder, C., & Watkins, C. L. (1990). Precision teaching and direct instruction: Measurably superior instructional technology in schools. *Performance Improvement Quarterly*, *3*, 74–96.

Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives: Vol.1. Cognitive domain*. New York: Longmans, Green.

Boreham, N. C. (1984). Personality factors related to self-reported lapse of attention during lectures. *Psychological Reports*, *55*, 76–78.

Boyce, T. E., & Hineline, P. N. (2002). Interteaching: A strategy for enhancing the user-friendliness of behavioral arrangements in the college classroom. *The Behavior Analyst*, *25*, 215–226.

Buskist, W., Cush, D., & DeGrandpre, R. J. (1991). The life and times of PSI. *Journal of Behavioral Education*, *1*, 215–234.

Buskist, W., & Saville, B. K. (2004). Rapport-building: Creating positive emotional contexts for enhancing teaching and learning. In B. Perlman, L. I. McCann, & S. H. McFadden (Eds.), *Lessons learned: Practical advice for the teaching of psychology* (Vol. 2, pp. 149–155). Washington, DC: American Psychological Society.

Buskist, W., Sikorski, J., Buckley, T., & Saville, B. (2002). Elements of master teaching. In S. F. Davis & W. Buskist (Eds.), *The teaching of psychology: Essays in honor of Wilbert J. McKeachie and Charles L. Brewer* (pp. 27–39). Mahwah, NJ: Erlbaum.

Cooper, J. O., Heron, T. E., & Heward, W. L. (1987). *Applied behavior analysis*. Columbus, OH: Merrill.

Dickson, K. L., Miller, M. D., & Devoley, M. S. (2005). Effect of textbook study guides on student performance in introductory psychology. *Teaching of Psychology*, *32*, 34–39.

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology*, *84*, 795–805.

Duch, B. J., Groh, S. E., & Allen, D. E. (Eds.). (2001). *The power of problem-based learning: A practical "how to" for teaching undergraduate courses in any discipline*. Sterling, VA: Stylus.

Engelmann, S., & Carnine, D. W. (1982). *Theory of instruction: Principles and application*. New York: Irvington.

Flora, S. R., & Logan, R. E. (1996). Using computerized study guides to increase performance on general psychology examinations: An experimental analysis. *Psychological Reports*, *79*, 235–241.

Fox, E. J. (2004). The personalized system of instruction: A flexible and effective approach to mastery learning. In D. J. Moran & R. W. Malott (Eds.), *Evidence-based educational methods: Advances from the behavioral sciences* (pp. 201–221). New York: Academic Press.

Griffin, M. M., & Griffin, B. W. (1998). An investigation of the effects of reciprocal peer tutoring on achievement, self-efficacy, and test anxiety. *Contemporary Educational Psychology*, *23*, 298–311.

Halpern, D. F. (2004). Creating cooperative learning environments. In B. Perlman, L. I. McCann, & S. H. McFadden (Eds.), *Lessons learned: Practical advice for the teaching of psychology* (Vol. 2, pp. 149–155). Washington, DC: American Psychological Society.

Hansen, T. L. (1990). A positive reinforcement program for controlling student absenteeism. *College Student Journal*, *24*, 307–312.

Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goals structures on achievement. *Psychological Bulletin*, *89*, 47–62.

Kazdin, A. E. (1982). *Single-case research designs: Methods for clinical and applied settings*. New York: Oxford University Press.

Keller, F. S. (1968). Good-bye teacher... *Journal of Applied Behavior Analysis*, *1*, 79–89.

Lamal, P. A. (1984). Interest in PSI across sixteen years. *Teaching of Psychology*, *11*, 237–238.

Lindsley, O. R. (1964). Direct measurement and prosthesis of retarded behavior. *Journal of Education*, *147*, 62–81.

Lloyd, K. E., Garlington, W. K., Lowry, D., Burgess, H., Euler, H. A., & Knowlton, W. R. (1972). A note on some reinforcing properties of university lectures. *Journal of Applied Behavior Analysis*, *5*, 151–155.

Lloyd, M. E., & Lloyd, K. E. (1986). Has lightning struck twice? Use of PSI in college classrooms. *Teaching of Psychology*, *13*, 149–151.

Mathie, V. A., Beins, B., Benjamin, L. T. Jr., Ewing, M. M., Hall, C. C., Henderson, B., et al. (1993). Promoting active learning in psychology courses. In T. V. McGovern (Ed.), *Handbook for enhancing undergraduate education in psychology* (pp. 183–214). Washington, DC: American Psychological Association.

Michael, J. (1991). Behavioral perspectives on college teaching. *The Behavior Analyst*, *14*, 229–239.

Milem, J. F., Berger, J. B., & Dey, E. L. (2000). Faculty time allocation: A study of change over twenty years. *Journal of Higher Education*, *71*, 454–475.

Moran, D. J., & Malott, R. W. (Eds.). (2004). *Evidence-based educational methods: Advances from the behavioral sciences*. New York: Academic Press.

Saville, B. K., Zinn, T. E., & Elliott, M. P. (2005). Interteaching vs. traditional methods of instruction: A preliminary analysis. *Teaching of Psychology*, *32*, 161–163.

Schaefer, G., Epting, K., Zinn, T., & Buskist, W. (2003). Student and faculty perceptions of effective teaching: A successful replication. *Teaching of Psychology*, *30*, 133–136.

Shimoff, E., & Catania, A. C. (2001). Effects of recording attendance on grades in introductory psychology. *Teaching of Psychology*, *28*, 192–195.

Skinner, B. F. (1954). The science of learning and the art of teaching. *Harvard Educational Review*, *24*, 86–97. (Reprinted in B. F. Skinner, *Cumulative Record*, definitive edition, pp. 179–191, Cambridge, MA: B. F. Skinner Foundation, 1999)

Skinner, B. F. (1968). *The technology of teaching*. New York: Meredith.

Wilson, J. H., & Taylor, K. W. (2001). Professor immediacy as behaviors associated with liking students. *Teaching of Psychology*, *28*, 136–138.

Wright, M. C., Assar, N., Kain, E. L., Kramer, L., Howery, C. B., McKinney, K., et al. (2004). Greedy institutions: The importance of institutional context for teaching in higher education. *Teaching Sociology*, *32*, 144–159.

Yoder, J. D., & Hochevar, C. M. (2005). Encouraging active learning can improve students' performance on examinations. *Teaching of Psychology*, *32*, 91–95.

Received March 31, 2005
Final acceptance October 17, 2005
Action Editor, Edward Daly