

Sketch

INTERTEACHING: BRINGING BEHAVIORAL EDUCATION INTO THE 21ST CENTURY

Bryan K. Saville, Tonya Lambert, and Stephen Robertson

James Madison University

In recent years, educators have lamented the current state of the American education system, citing a number of possible reasons for its purported downfall. One potential contributor to this negative state of affairs is the continued use of ineffective teaching methods. Interteaching, a new teaching method that has its roots in behavior analysis, offers an effective and user-friendly alternative to traditional approaches to classroom instruction. In this article, we describe interteaching. We also discuss a number of studies that have examined the efficacy of interteaching relative to more common teaching methods. Finally, we provide recommendations for researchers wishing to study this promising behavior-analytic teaching method.

Key words: interteaching, education, behavioral teaching methods, lecture, component analysis

Although it might be premature to suggest that the American education system is in a state of crisis, there is nevertheless concern among educators that failure to make significant changes soon might eventually result in detrimental consequences. In their edited volume *Declining by Degrees*, Hersh and Merrow (2005), for example, lamented the current state of education in the United States: "Higher education, long viewed as the crown jewel of American education, is tarnished" (p. 1). Hersh and Merrow then suggested that "there are few reliable measures of what students actually learn or are able to do *because* they attended a particular college or university" (p. 3, italics in original). Throughout this volume, esteemed educators proposed reasons why the American education system is failing, among them the prevailing belief that college is nothing more than job preparation, the inability of students (and, often, their teachers) to synthesize large amounts of discrepant information, insufficient preparation for college, grade inflation, the disconnection between what students want in a college and what colleges are offering, and the overwhelming emphasis on the bottom line (i.e., money).

To this list, we wish to add another potential culprit: ineffective teaching methods. Lecture, christened "the Velveeta of teaching methods" by Benjamin (2002, p. 57), is the pedagogical weapon of choice for more than 80% of college teachers. Enter any college classroom, and there is a good

chance one will observe a teacher, standing at the front of the class, imparting information—often via meticulously crafted PowerPoint slides, complete with every bell and whistle needed to capture one’s attention—to a room full of students, some of whom might be trying to jot down every word that comes out of the teacher’s mouth, others of whom might be dozing off, reading the paper, texting a friend, or staring off into space. Unfortunately, traditional lectures are rather ineffective at increasing student learning (e.g., McKeachie & Svinicki, 2006). One might therefore assume that psychology teachers, many of whom make their living studying the ways people learn best, would be likely to use other, more effective teaching methods. In reality, though, most psychology teachers have continued to use traditional methods that fail to capitalize on what psychological researchers have discovered about human learning. As Halpern and Hakel (2003) stated:

[There is] precious little evidence that content experts in the learning sciences actually apply the principles they teach in their own classrooms. Like virtually all college faculty, they teach the way they were taught. But, ironically (and embarrassingly), it would be difficult to design an educational model that is more at odds with the findings of current research about [human learning] than the one being used today at most colleges and universities. . . . Psychologists, educators, and other professionals already have available to them a substantial body of research that can be drawn upon to inform those responsible for designing and implementing learning programs. Unfortunately, the research literature is usually ignored, while educational leaders and policymakers grasp at the “ephemeral” magic of quick fixes. (pp. 37-38)

Over the past several decades, behavioral researchers have designed and implemented a number of alternative teaching methods, all of which have their roots in B. F. Skinner’s (1938, 1953) operant psychology. Although these methods tend to be more effective than traditional teaching methods, they, too, have gained little traction in college classrooms, often for reasons we will discuss shortly. The purpose of this article is to discuss a new behavior-analytic teaching method called *interteaching* (Boyce & Hineline, 2002). In short, we believe that interteaching might appeal to teachers who wish to shake things up a bit and who want to use evidence-based methods in their classrooms. Before introducing interteaching, however, we will briefly review other behavioral approaches to education and why some of these methods have waned in popularity in the last 20 to 30 years, despite evidence that they improve learning more than traditional teaching methods (e.g., Born, Gledhill, & Davis, 1972; Fernald & Jordan, 1991; Kulik, Kulik, & Cohen, 1979; Taveggia, 1976). We will also discuss the promising results from a number of studies that have examined the efficacy of interteaching. Finally, we will provide suggestions for future research on interteaching.

Behavioral Approaches to Education

Behavioral approaches to education appeared in the 1950s, when B. F. Skinner (1954/1999) first discussed the application of behavioral principles in the classroom. In his seminal paper, Skinner identified several problems

with traditional approaches to classroom instruction—problems that, for the most part, still pervade pedagogical practices. First, teachers rarely use positive reinforcement in any systematic way, choosing instead to use aversive consequences to modify behavior. As Skinner noted, the use of aversive consequences has produced students who study and attend class not because they enjoy doing so, but because they can avoid the aversive consequences (e.g., receiving a failing grade) that might otherwise occur. In addition, Skinner noted that when teachers do use positive reinforcement (a) the contingencies are often imperfect (e.g., reinforcement is delayed or uncertain), and (b) there is little emphasis on shaping, the result of which is target behaviors that are often slow to emerge.

In response to Skinner's (1954/1999) concerns, behavioral researchers developed several teaching methods—Lindsley's (1964) precision teaching, Skinner's (1968) programmed learning, Keller's (1968) personalized system of instruction (PSI), and Engelmann and Carnine's (1982) direct instruction, to name a few—that attempted to capitalize on well-known principles of learning and alleviate many of the problems inherent in more traditional teaching practices. Although these methods vary somewhat in format and scope, they share certain features (Kinder & Carnine, 1991; see also Moran & Malott, 2004). First, each program uses positive reinforcement, and not aversive consequences, to produce desired student-learning outcomes. Second, the programs typically focus on student mastery of course material. Third, student performance is assessed frequently. And finally, the presentation of course material moves from simple to complex, with complex concepts broken down into simpler subcomponents that students must master before moving on. Although extensive discussion of these teaching methods is beyond the scope of this article, suffice it to say that behavioral teaching methods have consistently outperformed more traditional teaching methods at all levels of education (for reviews, see Binder & Watkins, 1990; Buskist, Cush, & DeGrandpre, 1991; Jaehnig & Miller, 2007; Kinder & Carnine, 1991; Moran & Malott, 2004).

Why Are These Methods No Longer Used?

If behavioral teaching methods do much to improve student learning, why, then, do most college teachers continue to use other methods? There are several possible reasons. First, as Buskist et al. (1991) suggested, behavioral teaching methods often do not jibe with traditional pedagogical practices. PSI, for instance, with its focus on self-pacing, may require more than the usual 10- or 15-week format that is common in most educational settings. In addition, because many behavioral teaching methods focus on mastery, they typically produce high grades, an outcome that some educators view as grade inflation.

Second, habit, coupled with a lack of time, may be responsible for the relative indifference that many teachers have shown toward behavioral teaching methods (Buskist et al., 1991). Many teachers—including those who are aware of the positive outcomes generated by behavioral teaching methods—have continued to lecture simply because they have done so for years. Moreover, because most teachers have other activities that compete for their time (e.g., service activities, grant writing), many choose lecture-based methods because they tend to be less time-consuming than behavioral teaching

methods—some of which require extensive preparation time, at least initially (Boyce & Hinline, 2002).

Third, student resistance to alternative teaching methods in general may lead some teachers to use methods that students enjoy most or with which students are most familiar. By college, most students equate teaching with lecturing, and many have learned how to succeed, relatively speaking, under such conditions. Thus, many students have come to prefer lectures—regardless of whether they actually improve learning. Because student evaluations continue to be a source of information about teaching efficacy, and because teaching efficacy is one marker used in tenure and promotion decisions, some teachers prefer to use methods that net them the best student evaluations—again, regardless of whether they actually improve learning.

Fourth, some teachers are hesitant to give up control in their classrooms. With most behavioral teaching methods, focus shifts from teacher to students, who are largely responsible for their own learning. Consequently, teachers function as facilitators of learning rather than as omniscient dispensers of information. Teachers who enjoy their role as “sage on the stage” may dismiss behavioral methods because they see themselves as more than facilitators of learning.

Finally, misunderstanding of behavioral principles, and the underlying behavioral philosophy on which these principles stand, has likely contributed to the scarcity of behavioral teaching methods in college classrooms. Kohn (1993), for instance, in his popular book *Punished by Rewards*, advocated against the use of applied behavior analysis, or what he referred to as the “do this and you’ll get that” (p. 3) approach to behavior change. Sadly, several of Kohn’s assumptions—for instance, that reinforcement and reward are functionally equivalent and that reinforcement (a consequence) is akin to bribery (an antecedent)—are inaccurate. Similarly, Ainsworth (1979), who referred to PSI as “an innovation that failed” (p. 42), argued that self-paced instruction is not practical because it requires students to have high levels of motivation at the outset. Unfortunately, Kohn and Ainsworth—along with many other critics of behavior analysis (e.g., Brewer, 1974; Chomsky, 1959)—seem to have misunderstood how behavioral principles operate.

Description of Interteaching

In an attempt to overcome some of the resistance that educators have shown toward earlier behavioral teaching methods, Boyce and Hinline (2002) introduced interteaching, a user-friendly alterative that “retains some key characteristics of [earlier behavioral teaching methods] but offers greater flexibility for strategies that are based on behavioral principles” (p. 215). The typical format for interteaching proceeds as follows. First, the teacher constructs a preparation (prep) guide¹ consisting of questions designed to guide students through a reading assignment. The questions cover a range of formats, often proceeding from simpler definitional-type questions to more complex application and synthesis questions (thus, using a shaping-type format that Skinner, 1954/1999, noted was absent from most teaching methods; see also Bloom, 1956; Ray & Belden, 2007). Usually, a prep guide contains 10 to 12 items, each of which may consist of more than

1 For sample prep guides, see the Appendix of Boyce and Hinline (2002), or contact the first author of this article.

one question, and covers 10 to 15 pages of material, although that number may vary depending on a number of factors (e.g., complexity of the material, number of days between classes). The teacher then distributes the prep guide to students (e.g., via a course Web page), who then have several days to complete the prep-guide items before class. In class, students first hear a brief clarifying lecture that reviews selected material from the previous class period (see below). After the lecture, students form pairs to discuss the prep guide. Although students are free to work through the prep guide at their own pace, speedy completion is likely a sign that students have not discussed the material thoroughly. If students discuss the material thoroughly, the pair discussions should last approximately two thirds of the class period (e.g., 50 min in a 75-min class).² During the pair discussions, the teacher (and teaching assistant, if available) traverses the classroom, answering questions and guiding the discussions. After students have discussed the prep guide thoroughly, they complete a record sheet, which provides the teacher with feedback on how the discussions went and which material was difficult to understand.

Based on comments on the record sheets (along with information the teacher may have gleaned while interacting with students), the teacher prepares a clarifying lecture that begins the next class period. The clarifying lecture, which lasts approximately one third of the class period, covers the prep-guide items most requested by students,³ along with any supplementary material the teacher might wish to discuss. Because students dictate which material the teacher will cover during the lectures, and because the lectures *follow* the behaviors of interest (e.g., reading before class, discussing the material), the clarifying lectures are more likely to function as consequences for desired behaviors and not as antecedents, which tends to be the case with lecture-based courses. After the clarifying lecture is complete, students spend the rest of the class period discussing the next prep guide.

There are additional components to interteaching as well (Boyce & Hinline, 2002). First, as with other behavioral teaching methods, testing occurs frequently—at least five times per semester. Frequent testing is useful for a number of reasons: (a) Students have ample opportunity to show what they have learned; (b) course grades are not greatly affected if students take awhile to get the hang of interteaching or if other factors preclude them from doing well on every exam (for this reason, Boyce & Hinline, 2002, proposed dropping each student's lowest exam grade); and (c) frequent testing produces spaced practice (Donovan & Radosevich, 1999) and may capitalize on the testing effect, a phenomenon in which test taking improves long-term retention

2 We have found that students sometimes do not need two thirds of a class period to complete their discussions thoroughly (e.g., when certain material is fairly easy to comprehend). Although one might be tempted in this case simply to add items to the prep guide, we have found that prep guides much longer than 12 or so items become aversive to students. To resolve this problem, we often distribute in class additional items that students must discuss after they have completed the prep guides. This accomplishes the goals of giving students additional practice and increasing discussion time without increasing the length of the prep guides.

3 We usually tell students that we will go over no more than three or four prep-guide items during the clarifying lectures. As a result, students seem to ask more questions while in class, rather than hoping we will cover every item during the next lecture; this also allows us to introduce supplementary material without greatly increasing the length of the lectures.

of information (Roediger & Karpicke, 2006). The exam content should be tied to the prep-guide material so that there is a clear link between the behaviors students practice during the in-class discussions and the behaviors they emit while taking exams. Each exam should have one or two essay questions (these questions are usually conceptual in nature and taken directly from the prep guides) and a number of objective questions that cover the remaining prep-guide material. In this way, when students take exams, the questions should exert strong stimulus control over their behavior.

Second, for participating in each pair discussion, students receive a small number of points that eventually total 10% of their course grade. In general, students who attend class get better grades than students who miss class (e.g., Shimoff & Catania, 2001). Thus, participation points may motivate students to attend class, which may improve course performance and, ultimately, long-term learning.

Third, interteaching includes quality points, a component designed to improve the quality of students' in-class discussions by introducing an explicit cooperative contingency (i.e., the outcomes of one person partially depend on the performance of another). Imagine two students, Steve and Erin, who last week discussed a prep guide together and subsequently took an exam that contained one of the prep-guide items as an essay question. If Steve and Erin each get an A or B on that essay question (e.g., four or five points on a five-point question), they receive a small number of "quality" points toward their overall course grade. If, however, either Steve or Erin receives a C or lower on that question (i.e., three or fewer points), neither gets quality points. Boyce and Hineline (2002) suggested that quality points should account for approximately 10% of students' overall course grades.

Is Interteaching Effective?

In their original article, Boyce and Hineline (2002) provided anecdotal evidence that interteaching might be an effective teaching method; they did not, however, provide any systematic evidence of its efficacy. Saville, Zinn, and Elliott (2005) conducted the first systematic investigation of interteaching, a lab-based study that compared interteaching to more traditional teaching methods. They randomly assigned students to one of four conditions—interteaching, lecture, reading, or control—and exposed students in the first three conditions to material from an article on pet ownership and health (Allen, 2003). Students then returned to the lab 1 week later and took a 10-item, multiple-choice quiz. Saville et al. found that students in the interteaching condition (74%) answered a greater percentage of questions correctly than students in the other three conditions. Moreover, the average scores for students in the lecture (60%), reading (55%), and control (52%) conditions were not significantly different from one another. Saville et al. also discussed several limitations to their study, including the fact that multiple-choice quizzes may not be best for identifying differences in learning and that the article they used (Allen, 2003) contained information on a topic (i.e., the relation between pet ownership and health) about which students may have been knowledgeable before participating in the study (hence, one possible reason why students in the control condition scored above chance on the quiz). Nevertheless, because students in the interteaching condition performed significantly better than students in the lecture and readings

conditions—who performed no better than students in a control condition after just 1 week—Saville et al. concluded that interteaching might be an effective alternative to other teaching methods.

To examine the generality of Saville et al.'s (2005) findings, Saville, Zinn, Neef, Van Norman, and Ferreri (2006) compared interteaching to lecture in two different college courses. In Study 1, students in a graduate-level special education course took pretests at the beginning of the quarter and then alternated between interteaching and lecture sessions over the next 8 weeks. Although posttest quiz scores were higher than pretest scores under both teaching conditions, the difference was consistently greater with interteaching. In addition, most students preferred interteaching to lecture. In Study 2, students in two sections of an undergraduate research-methods course alternated several times between interteaching and lecture, the order of which was counterbalanced across sections (i.e., for each unit of material, one section participated in interteaching while the other section heard lectures). After each unit of material, students in both sections took the same exam. Students in the interteaching condition consistently received higher exam grades than students in the lecture condition. And as in Study 1, the majority of students in both sections preferred interteaching.

Researchers have also begun to conduct component analyses of interteaching. Saville and Zinn (2009) examined the quality-points component of interteaching by comparing the exam scores of students in two sections of an introductory psychology course. Saville and Zinn alternated between interteaching with quality points and interteaching without quality points, counterbalanced across sections, and awarded quality points by comparing the essay answers of students who completed the pair discussions together (see Boyce & Hinline, 2002). Saville and Zinn found that the average exam grades in the two sections were not significantly different on five of six exams. They suggested that the delivery of quality points was too delayed to have an effect on exam scores—students often did not receive their exam grades until a week or more after completing the discussions—and that more immediate consequences (e.g., social consequences delivered by discussion partners and the teacher) likely had a greater impact on learning.

Finally, Goto and Schneider (2009) examined student responses to interteaching in two upper-level nutrition courses. They found that students enjoyed interteaching and reported that it fostered critical thinking. Although Goto and Schneider did not systematically compare student-learning outcomes produced by interteaching and other teaching methods, they concluded that interteaching might be an effective and enjoyable way to teach nutrition courses.

In sum, although there are but a few published studies on interteaching, the results of these studies are promising: Interteaching seems to be an effective alternative to lecture, the most common form of classroom instruction, and students, for the most part, seem to enjoy interteaching. Nevertheless, there is still much to be done with this new and potentially promising teaching method.

Future Directions

As noted previously, a growing body of research suggests that interteaching might be an effective and enjoyable alternative to more traditional

pedagogical arrangements (Goto & Schneider, 2009; Saville et al., 2005, 2006). As with any new area of research, though, there is still extensive room for examination and a vast set of questions that have yet to be answered. In this section, we provide some recommendations regarding the future of interteaching research. The list is by no means exhaustive; in fact, given the paucity of interteaching research conducted thus far, we assume that other researchers will inevitably identify important questions that we have yet even to ponder.

Direct and Systematic Replication

Saville et al.'s (2005, 2006) findings provide promising evidence regarding the efficacy of interteaching. Nevertheless, as Sidman (1960) pointed out nearly a half-century ago, direct replication is vital to establishing the validity of an observation. Thus, to confirm that interteaching does produce superior student-learning outcomes relative to more traditional teaching methods, researchers will need to replicate these earlier studies, conducting both lab-based and classroom-based analyses that examine the efficacy of interteaching. Sidman also discussed the importance of systematic replication, in which a researcher examines whether a phenomenon occurs under conditions that differ from those prevailing in the original experiment. Although we have used interteaching in a number of our undergraduate and graduate courses, the researchers who have thus far published their findings on interteaching have used relatively small convenience samples consisting of introductory psychology students (see Saville et al., 2005; Saville & Zinn, 2009), students in undergraduate research methods courses (Saville et al., 2006, Study 2), or students in a graduate-level special education course (Saville et al., 2006, Study 1). It would therefore be interesting to examine whether interteaching produces similar results in other types of psychology courses—large, survey-level courses; smaller, capstone-level courses; graduate-level seminars; and even distance-learning courses. Researchers would also do well to examine interteaching in nonpsychology courses to determine whether interteaching has a positive effect on learning in these courses as well. There is no reason to assume that interteaching would not be effective in other types of courses, and Goto and Schneider (2009) have initiated this endeavor by examining the social validity of interteaching in two upper-level nutrition courses.

Component Analyses

Just as research helped determine which components were important to the success of PSI (for reviews, see Buskist et al., 2001; Fox 2004), component analyses will be vital to understanding what makes interteaching a potentially effective teaching method. As mentioned previously, Saville and Zinn (2009) examined the quality-points component of interteaching by giving points contingent on essay performance. They found that systematically manipulating quality points across two sections of an introductory psychology course did not have an effect on exam scores. Nevertheless, because cooperative contingencies have positive effects on various performance measures (e.g., Johnson, Maruyama, Johnson, Nelson, & Skon, 1981), researchers might wish to examine other ways of manipulating quality points. Saville

and Zinn, for instance, suggested that increasing the number of quality points that students receive might have an effect on performance. Similarly, Saville et al. (2006, Study 2) had teaching assistants distribute tickets to students who were engaged in quality discussions; the tickets were then entered into a drawing in which the students had a chance to win bonus points. Saville et al. did not, however, examine whether this manipulation affected exam performance. Finally, teachers could use performance on short post-discussion quizzes rather than exam performance as a source of quality points. This contingency would be similar to the one described by Boyce and Hinline (2002) and tested by Saville and Zinn, but it would provide more immediate consequences for students.

Another component of interteaching ripe for examination is the prep-guide component. Because the prep guides serve as a roadmap for learning, it seems likely that they might prove to be a necessary component of interteaching. Comparing interteaching with prep guides to interteaching without prep guides (i.e., students simply read the chapters before class and then discuss the material during pair discussions) would help determine the role that prep guides play in the interteaching method. With regard to prep-guide content, Boyce and Hinline (2002) recommended that prep guides include a blend of factual, definitional, and analytical questions. Because the intended purpose of the prep guides is to shape students' behavioral repertoires, manipulating the types of questions included would help determine how each contributes to learning. Such a study might also help teachers construct better prep guides.

The pair discussions comprise the core classroom experience of interteaching, providing students with the opportunity to teach the course material to one another and practice the skills that define learning. A good amount of research shows that getting students actively involved tends to have positive effects on learning (e.g., Mathie et al., 1993). As such, it would be worthwhile to examine the extent to which the pair discussions contribute to the success of interteaching. At times, some of our students have wondered why they have to participate in pair discussions if they have already completed the prep-guide items and can get clarification during the lectures. Studies comparing interteaching with pair discussions to interteaching without pair discussions (i.e., students complete the prep guides, submit a record sheet, and hear a clarifying lecture) would provide teachers with an answer to this question.

Making Interteaching More Effective

Although interteaching seems to have positive effects on learning relative to more traditional teaching methods, there are likely other ways to make interteaching even more effective. In their original description of interteaching, Boyce and Hinline (2002) recommended frequent testing. They suggested, however, that teachers should avoid daily quizzes as means of controlling student behavior because the quizzes, especially when unannounced, might take on aversive properties (although see Connor-Greene, 2002, for a study in which students preferred daily quizzes over less frequent testing). Recent research on the testing effect, though, has shown that testing—even in the absence of feedback—seems to improve long-term retention of material (e.g., McDaniel, Anderson, Derbish, & Morrisette,

2007; Roediger & Karpicke, 2006). Given the positive effect that frequent testing seems to have on student learning, researchers might wish to examine whether more frequent—even daily—testing improves interteaching. This practice might capitalize on the testing effect and, as mentioned earlier, might also be an interesting way of improving the quality of pair discussions.

The recent invention of personal response systems, or “clickers,” provides another interesting avenue of research for those interested in studying interteaching. Clickers are electronic handheld devices that allow students to answer multiple-choice questions projected on a screen. Once students answer a question, a distribution of responses appears, and the teacher can then discuss the correct answer. A growing number of studies suggest that the use of clickers in college classrooms has a significant impact on learning because students are actively involved during lectures and receive immediate clarification for any misunderstandings they might have (Hoekstra, 2008; Mayer et al., 2009; Morling, McAuliffe, Cohen, & DiLorenzo, 2008). With interteaching, students receive clarification on difficult questions during lectures that usually begin the subsequent class period. In most cases, though, the information presented in the clarifying lectures is delayed, often by 2 or more days, and thus might have a diminished effect on learning (see, e.g., Brosvic & Epstein, 2007). Using clickers might be one way to minimize this delay. For instance, once teachers have an idea of which prep-guide items might cause difficulty, they could interrupt the pair discussions for a few minutes and ask students to respond to one or two multiple-choice items. If students seem to be having trouble with a particular concept or problem, the teacher could take some time to provide clarification. In this way, much of the clarifying lecture would get intertwined with the pair discussions rather than presented at the beginning of the next class period.

Again, these are just some of the numerous possibilities that exist for researchers interested in studying interteaching. We hope that researchers will continue to examine some of these questions along with others that pique their interest.

Conclusion

The behavioral principles underlying interteaching are well-established and have been applied successfully in many settings, with many populations. There is no reason to believe that the application of behavioral principles in the form of interteaching would fail to produce positive behavior change. Although the extant research is promising, additional research is nevertheless needed to validate interteaching as a new behavioral technology for education. With continued examination and application, we believe that interteaching has the potential to be an effective and enjoyable alternative to traditional teaching methods, one that helps lead behavioral education into the 21st century.

References

- AINSWORTH, L. L. (1979). Self-paced instruction: An innovation that failed. *Teaching of Psychology*, 6, 42-46.

- ALLEN, K. (2003). Are pets a healthy pleasure? The influence of pets on blood pressure. *Current Directions in Psychological Science*, 12, 236-239.
- 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.
- BORN, D. G., GLEDHILL, S. M., & DAVIS, M. L. (1972). Examination performance in lecture-discussion and personalized instruction courses. *Journal of Applied Behavior Analysis*, 5, 33-43.
- 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.
- BREWER, W. F. (1974). There is no convincing evidence for operant or classical conditioning in adult humans. In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes* (pp. 1-42). Hillsdale, NJ: Erlbaum.
- BROSVIC, G. M., & EPSTEIN, M. L. (2007). Enhancing learning in the introductory course. *The Psychological Record*, 57, 391-408.
- BUSKIST, W., CUSH, D., & DEGRANDPRE, R. J. (1991). The life and times of PSI. *Journal of Behavioral Education*, 1, 215-234.
- CHOMSKY, N. (1959). Review of Skinner's *Verbal Behavior*. *Language*, 35, 26-58.
- CONNOR-GREENE, P. (2000). Assessing and promoting student learning: Blurring the line Between teaching and testing. *Teaching of Psychology*, 27, 84-88.
- 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.
- ENGELMANN, S., & CARNINE, D. W. (1982). *Theory of instruction: Principles and application*. New York: Irvington.
- FERNALD, P. S., & JORDAN, E. A. (1991). Programmed instruction versus standard text in introductory psychology. *Teaching of Psychology*, 18, 205-211.
- 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.
- GOTO, K., & SCHNEIDER, J. (2009). Interteaching: An innovative approach to facilitate university student learning in the field of nutrition. *Journal of Nutrition Education and Behavior*, 41, 303-304.
- HALPERN, D. F., & HAKEL, M. D. (2003). Applying the science of learning to the university and beyond. *Change*, 35(4), 36-42.
- HERSH, R. H., & MERROW, J. (Eds.). (2005). *Declining by degrees*. New York: Palgrave MacMillan.
- HOEKSTRA, A. (2008). Vibrant student voices: Exploring effects of the use of clickers in large college courses. *Learning, Media, and Technology*, 33, 329-341.
- JAETHNIG, W., & MILLER, M. L. (2007). Feedback types in programmed instruction: A systematic review. *The Psychological Record*, 57, 219-232.

- JOHNSON, D. W., MARUYAMA, G., JOHNSON, R., NELSON, D., & SKON, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement. *Psychological Bulletin*, *89*, 47-62.
- KELLER, F. S. (1968). Good-bye teacher *Journal of Applied Behavior Analysis*, *1*, 79-89.
- KINDER, D., & CARNINE, D. (1991). Direct instruction: What it is and what it is becoming. *Journal of Behavioral Education*, *1*, 193-213.
- KOHN, A. (1993). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes*. Boston: Houghton Mifflin.
- KULIK, J. A., KULIK, C. C., & COHEN, P. A. (1979). A meta-analysis of outcome studies of Keller's personalized system of instruction. *American Psychologist*, *34*, 307-318.
- LINDSLEY, O. R. (1964). Direct measurement and prosthesis of retarded behavior. *Journal of Education*, *147*, 62-81.
- 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.
- MAYER, R. E., STULL, A., DELEEUW, K., ALMEROOTH, K., BIMBER, B., CHUN, D., ET AL. (2009). Clickers in college classrooms. Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology*, *34*, 51-57.
- MCDANIEL, M. A., ANDERSON, J. L., DERBISH, M. H., & MORRISETTE, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, *19*, 494-513.
- MCKEACHIE, W. J., & SVINICKI, M. (2006). *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers* (12th ed.). Boston: Houghton Mifflin.
- MORAN, D. J., & MALOTT, R. W. (Eds.). (2004). *Evidence-based educational methods: Advances from the behavioral sciences*. New York: Academic Press.
- MORLING, B., MCAULIFFE, M., COHEN, L., & DILORENZO, T. M. (2008). Efficacy of personal response systems ("clickers") in large, introductory psychology classes. *Teaching of Psychology*, *35*, 45-50.
- RAY, R. D., & BELDEN, N. (2007). Teaching college level content and reading comprehensions skills simultaneously via an artificially intelligent adaptive computerized instructional system. *The Psychological Record*, *57*, 201-218.
- ROEDIGER, H. L., III, & KARPICKE, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249-255.
- SAVILLE, B. K., ZINN, T. E., & ELLIOTT, M. P. (2005). Interteaching versus traditional methods of instruction: A preliminary analysis. *Teaching of Psychology*, *32*, 161-163.
- SAVILLE, B. K., ZINN, T. E., NEEF, N. A., VAN NORMAN, R., & FERRERI, S. J. (2006). A comparison of interteaching and lecture in the college classroom. *Journal of Applied Behavior Analysis*, *39*, 49-61.
- SAVILLE, B. K., & ZINN, T. E. (2009). Interteaching: The effects of quality points on exam scores. *Journal of Applied Behavior Analysis*, *42*, 369-374.
- SHIMOFF, E., & CATANIA, A. C. (2001). Effects of recording attendance on grades in introductory psychology. *Teaching of Psychology*, *28*, 192-195.

- SIDMAN, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. Boston, MA: Authors Cooperative.
- SKINNER, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century-Crofts.
- SKINNER, B. F. (1953). *Science and human behavior*. New York: The Free Press.
- SKINNER, B. F. (1999). The science of learning and the art of teaching. In *Cumulative record, definitive edition* (pp. 179–191). Cambridge, MA: B. F. Skinner Foundation. (Original work published 1954)
- SKINNER, B. F. (1968). *The technology of teaching*. New York: Meredith Corporation.
- TAVEGGIA, T. C. (1976). Personalized instruction: A summary of comparative research, 1967–1974. *American Journal of Physics*, 44, 1028–1033.

