

## **Enhancing the impact of quality points in interteaching**

**Rocío Rosales<sup>12</sup>, James L. Soldner<sup>3</sup>, and William Crimando<sup>4</sup>**

*Abstract: Interteaching is a classroom instruction approach based on behavioral principles that offers increased flexibility to instructors. There are several components of interteaching that may contribute to its demonstrated efficacy. In a prior analysis of one of these components, the quality points contingency, no significant difference was reported in student exam scores when quality points were made available. The purpose of the present study was to further evaluate the impact of the quality points component of interteaching, and to enhance the immediacy of feedback provided to students on this contingency via delivery of an answer key upon submission of post-discussion quizzes with the opportunity to review the answer key with their interteaching partner during class. We hypothesized that student quiz scores would be higher when this quality points contingency was in place. An alternating treatments design was employed to compare student performance on post-discussion quizzes during two conditions: quality points vs. no quality points with all other components of interteaching in effect. Eleven undergraduate students enrolled in an introduction to applied behavior analysis course served as participants. Results indicate average quiz performance was higher following class sessions with the quality points contingency. Discussion is focused on the implications of delivering immediate feedback to students during classroom instruction and future directions of interteaching research.*

*Keywords: interteaching, quality points, feedback, college teaching*

Interteaching is a deviation from traditional lecture, with one of its main components, the “interteach” defined by Boyce and Hineline (2002) as a “mutually probing, mutually informing conversation...” (p. 220). Components of interteaching include the following: (1) preparation guides or “prep guides” which consist of 10-12 questions that cover roughly 10-15 pages of reading material, (2) pair discussions, conducted in class following the completion of each prep guide, (3) record sheets completed by students to provide feedback to the instructor and rate the overall quality of the pair discussion, (4) clarifying lectures designed to cover the most challenging topics based on feedback provided on the record sheets, (5) frequent test probes based on material from the prep guides to assess student learning, and (6) quality points, an explicit cooperative contingency whereby additional points are added to a student’s grade only if both students in a pair perform to a certain pre-determined criterion on a selected test probe.

Empirical support for the efficacy of interteaching in the college classroom, when compared to the traditional lecture, has been reported in previous studies. For example, Saville, Zinn, Neef, Van Norman, and Ferreri (2006) conducted two experiments to compare student

---

<sup>1</sup> Department of Psychology, University of Massachusetts Lowell, 113 Wilder Street, Suite 300, Lowell MA 01854-3059, [rocio\\_rosales@uml.edu](mailto:rocio_rosales@uml.edu)

<sup>2</sup> Author’s Note: The first and second author’s contributions should be considered equal.

<sup>3</sup> School for Global Inclusion and Social Development, University of Massachusetts Boston, 100 Morrissey Boulevard, Boston, MA 02125

<sup>4</sup> Rehabilitation Institute, Southern Illinois University Carbondale, Rehn Hall, Mail Code 4609, Carbondale, IL 62901

performance on short-answer quizzes following interteaching or traditional lecture. A total of 60 undergraduate students were recruited across the two studies. Results showed consistently higher quiz scores following interteaching sessions when compared to lecture. Arntzen and Hoium (2010) reported similar results with a group of undergraduate students that self-rated their level of understanding and knowledge when interteaching was implemented compared to traditional lecture. Overall, studies that have integrated interteaching into the classroom have consistently reported high academic achievement on homework and participation (Filipiak, Rehfeldt, Heal, & Baker, 2010; Saville, Pope, Lovaas, & Williams, 2012); or quiz and end-of-semester grades (Cannella-Malone, Axe, & Parker, 2009; Lambert & Saville, 2012; Saville, et al., 2006; Saville, Pope, Truelove, & Williams, 2012).

To date, only two published studies have included component analyses to identify the relative impact of each element of interteaching. For example, Saville, Cox, O'Brien, and Vanderbelt (2011) evaluated the impact of the clarifying lecture on the efficacy of interteaching. Results of this study indicated that students who received lectures had consistently higher exam scores, but these differences were only statistically significant on two exams delivered throughout the semester. Saville and Zinn (2009) examined the extent to which quality points influenced test scores in an undergraduate research methods course. Quality points are intended to enhance the value of pair discussions, overall student learning, and retention of material. When interteaching is incorporated, quality points are designed as a cooperative contingency to earn additional points (e.g., 8-10% of a final grade) if both students in a dyad score at or above a pre-determined level on test probes. Saville and Zinn (2009) implemented this contingency on a randomly assigned counterbalanced schedule across two sections of the same course. Three quality points were awarded if both students scored at least 80% on a pre-selected essay question for each exam. Results of this study indicated no statistically significant differences in exam scores.

One potential factor impacting these results may be the delay in the delivery of feedback. Specifically, students did not learn if quality points had been awarded until at least one week after completion of the exam. Prior research has shown that the immediacy of feedback is an important indicator of performance across a variety of behaviors and settings (Alvero, Bucklin, & Austin, 2001; Coddling, Feinberg, Dunn, & Pace, 2005; Renner, 1964). For example, Solomon, Klein, and Politylo (2012) conducted a meta-analysis of the single-case literature and found that immediate feedback was effective at increasing teachers' treatment integrity when a new intervention was introduced in the classroom. As additional evidence, Scheeler, McAfee, Ruhl, and Lee (2006) found that the use of corrective, immediate feedback delivered via wireless technology was an effective way to increase completion of three-term contingency trials by teachers. These findings are consistent with past research that has shown a negative correlation between a delay of a contingency and its impact on performance (Lattal, 1993). As a general rule: the sooner the delivery of feedback following desired behavior, the better (Daniels & Bailey, 2014). The immediacy of feedback may also impact student academic performance in the classroom.

Thus, the purpose of the present study was to further evaluate the impact of the quality points component of interteaching; and to enhance the immediacy of feedback provided to students on this contingency via delivery of an answer key immediately upon submission of post-discussion quizzes with the opportunity to review the answer key with their interteaching partner.

## Method

### Participants and Materials

Eleven undergraduate students (1 male, 10 female;  $M$  age = 21) enrolled in one section of an introductory course to applied behavior analysis served as participants. Students were of sophomore (3), junior (4) or senior (4) standing. Students were not informed of the purpose of the study until the last day of the semester, at which time they were asked to review and sign a consent form that allowed use of their data. Students were given the assurance that the instructor would not review the consent forms until final grades were submitted for the term.

The class met face-to-face on Mondays and Wednesdays for 75 min. and was web-enhanced (i.e., incorporated the use of *Blackboard Learn* to distribute materials to the students throughout the semester). The textbook adopted for the course was *Principles of Everyday Behavior Analysis* (Miller, 2006). Preparation guides, created by the instructor for the course (second author), were made available to students at least one week before they were expected to discuss the material with a classmate during class time. The prep guides reviewed 5 to 10 pages of material and included definition, conceptual, and applied questions (see Appendix A for an example of questions included in the prep guides).

### Dependent Measure and Experimental Design

An alternating treatments design was implemented whereby a quality points contingency was in effect via quasi-random assignment throughout the semester (quality points were not available during every class meeting). The availability of quality points was determined prior to the start of the semester by a coin flip with the constraint that each condition (quality points vs. no quality points) could occur for no more than two consecutive sessions. Students were informed of the condition in place at the beginning of each pair discussion. That is, students were not aware if the contingency was in place until they began a pair discussion in class. The primary dependent variable for the study was average performance on 10-pt. weekly quizzes.

Quiz questions were based on the information presented in each prep guide and were created by adopting questions included as part of the textbook resources (see Appendix B for a sample quiz). Questions on each quiz required students to provide or apply a definition, recall information from the readings, or apply knowledge to novel examples. Question format included fill-in-the-blank and short-answer. The short answer questions were objective and based on examples provided in the instructor resources of the text (e.g., correct responses consisted of writing numbers or 1-2 word verbatim responses from the text, see Appendix B). A total of 17 quizzes were included as part of the data analysis, with a total of 34 possible quality points available throughout the semester. These points were considered extra credit and were added to the students' overall grade at the end of the term. Overall grades were determined by accumulating total quiz points and quality points, as well as various other required assignments (i.e., research article critique and book summaries) with objective point totals.

### Procedure

The general procedure for this study was similar to that reported by Saville and Zinn (2009). Pair discussions were held for each prep guide followed by a post-discussion quiz; record

sheets were completed following each pair discussion; and a clarifying lecture was delivered based on feedback provided by the students during the next class session. Students were allowed to select and work with one partner of their choosing, but were instructed to work with a different classmate during each pair discussion. If an odd number of students was present in class on any given day, one group was comprised of three students.

During the time allotted for discussion (25-30 min.), the instructor walked around the room to answer questions and monitor discussions to ensure students stayed on topic. Students received participation points for being present and prepared to discuss their prep guide with a classmate. Participation points comprised 8% of the students' final grade. Following each pair discussion, students were provided with approximately 5 min. to complete a record sheet intended to provide feedback to the instructor on the quality of the discussion, and to list the topics that presented the most difficulty. Quizzes were administered following completion of the record sheets. The instructor used the information from the record sheets to create a clarifying lecture that was presented at the beginning of the next class session and lasted approximately 25 min.

The cooperative contingency in place for the post-discussion quizzes in this study was as follows: If both students in a dyad received a quiz score of 80% (i.e., eight out of ten correct responses) or higher on their respective quiz, then two points were awarded for that particular quiz to each student. If either student received a score below 80%, then neither received quality points for that quiz. Quality points were calculated into the students' overall grade as bonus points. No other extra credit point opportunities were made available as part of the course. The quality points contingency was described in the course syllabus and the instructor reviewed this with the students at the beginning of the semester.

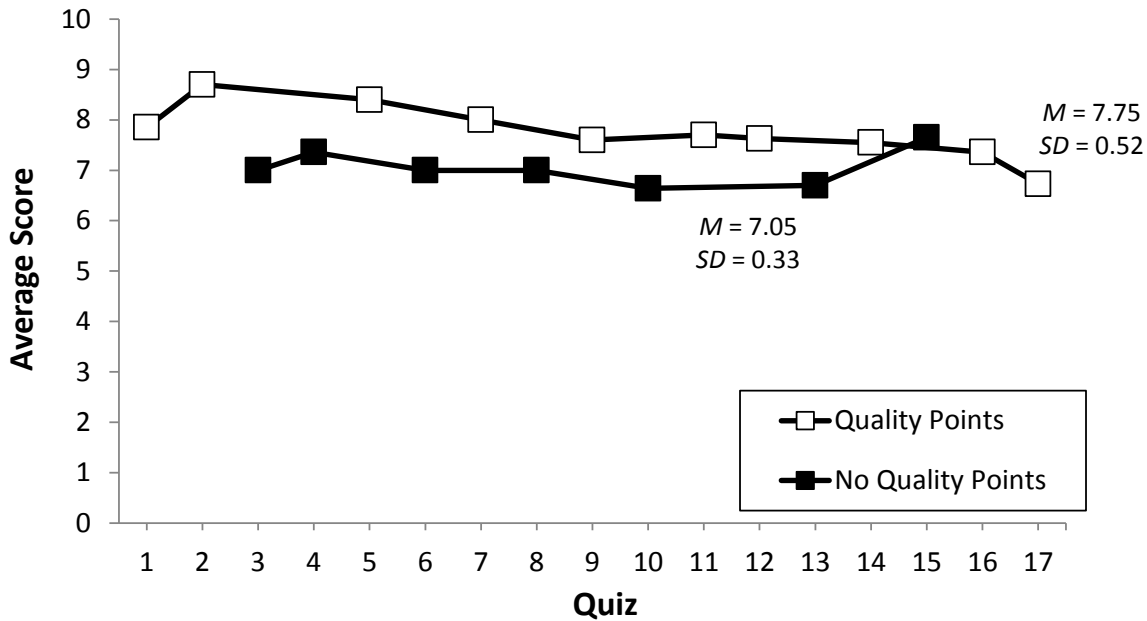
Immediate feedback was provided on the students' performance following every quiz regardless of the contingency in place. Specifically, the instructor provided the student with an answer sheet upon submission of each quiz. Upon receipt of the answer key, students had immediate information on their own performance and also had the opportunity to meet with their discussion partner (either in the classroom or in the hallway outside of the classroom) to discuss individual quiz performance and determine if the contingency for that respective quiz had been met.

## Results

The data analysis includes only scores for quizzes that were administered during class time immediately following pair discussions as described above. That is, if students missed class and subsequently received a score of '0' for the quiz, these grades were omitted for the purpose of the present study. If students missed class, but took a make-up quiz at a later time, these scores were also omitted. The quality points contingency was in effect for a total of 10 in-class quizzes administered throughout the semester, while the absence of quality points was in effect during seven quizzes.

Results of average quiz performance for the 11 students in the course are depicted in Figure 1. These results indicate overall average student performance was better when the cooperative contingency was in effect ( $M = 7.75$ ,  $SD = 0.52$  with quality points, compared to  $M = 7.05$ ,  $SD = 0.33$  with no quality points available). Overall, average scores were indicative of a difference in letter grade (e.g., 'C+' with quality points contingency compared to 'C-' without quality points contingency). Overall student performance tended to decrease as the semester

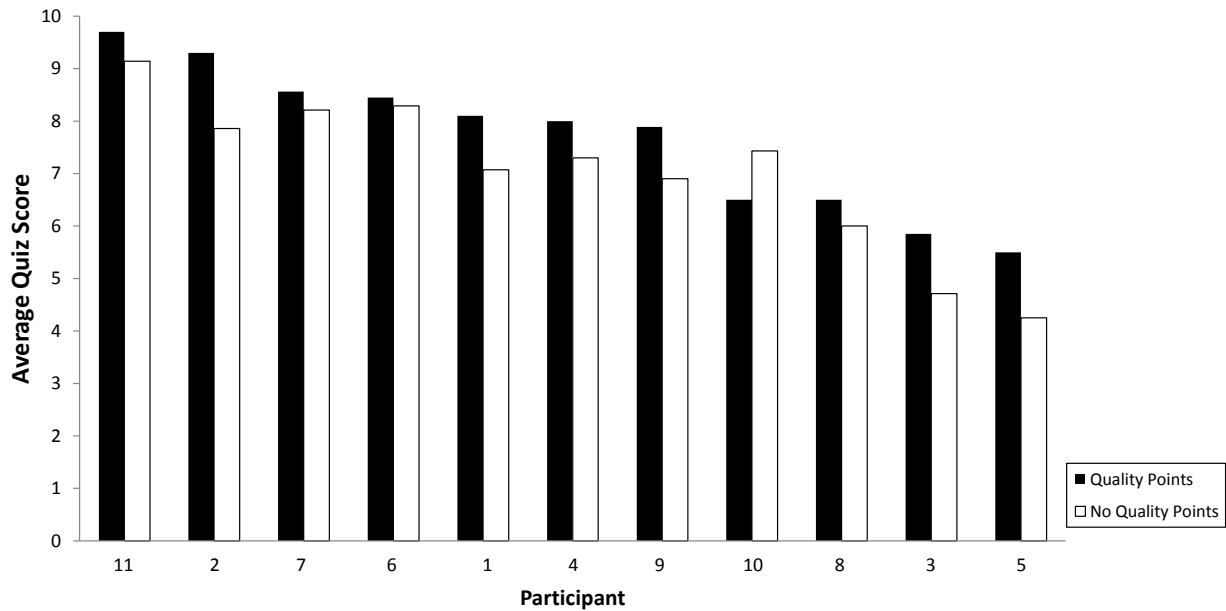
progressed. This may be accounted for by the fact that the information presented in the course was cumulative in nature. That is, after the basic principles of behavior analysis were presented and discussed, students were required to have a general understanding of these principles in order to perform well on subsequent chapter probes (e.g., differential reinforcement of behavior or shaping). Although the quizzes were not cumulative per se, the information covered in each subsequent chapter may have required students to retain information from previous chapters, resulting in slightly more difficult quizzes as the semester progressed.



**Figure 1.** Average performance across all post-discussion quizzes.

To investigate whether the differences between students' performance under the quality points contingency and absent the contingency were significant, a paired t-test was computed, using the means of all students under both conditions as scores ( $M = 7.75$ , with quality points,  $M = 7.05$  with no quality points). Since means comparisons are sensitive to extreme outliers, mean differences were converted to z scores in a preliminary analysis and then examined to determine if any of them exceeded 3 standard deviations either side of the grand mean. The largest z score was  $-2.4$ , indicating the absence of extreme outliers. Thus we proceeded with the paired t test. The difference was significant,  $t_{10} = 3.29, p = .008$ .

Individual average quiz scores are presented in Figure 2. These results reveal that 10 of the 11 participants scored higher on quizzes when quality points were available. The distribution of final course grades resulted in 6 students (55%) with a final course grade of A, 3 students (27%) with a final course grade of B, 1 student (9%) with a final course grade of C, and 1 student (9%) with a final course grade of D.



**Figure 2.** Individual mean scores across participants for quality vs. no-quality points condition.

### Discussion

The present results provide support for the use of quality points when interteaching is implemented in the classroom. Although the term ‘quality points’ was coined by Boyce and Hineline (2001), the use of a cooperative contingencies have been applied in classrooms through the use of strategies such as reciprocal peer tutoring (Bowman-Perrot et al., 2013) and other related paradigms.

Students in the present study received immediate feedback on all quizzes. Previous research evaluating the overall effectiveness of performance feedback has revealed that delayed feedback has less impact on performance than immediate feedback (Coddling, Feinberg, Dunn, & Pace, 2005; Daniels & Bailey, 2014; Lattal, 1993). The immediacy of feedback in the present study may have enhanced the effect of quality points on quiz performance and subsequently, student learning. Alternatively, immediate feedback may have enhanced performance *throughout* the semester, and not just when quality points were available. Although results indicate average quiz scores were higher when the quality points contingency was in effect, these results should be interpreted with caution given some of the limitations of the study.

First, the availability of quality points was not counterbalanced with a second class section. Therefore, it is possible that the content of the quizzes was not equal across sessions. This presents a potential extraneous variable since students may have found material more difficult in class sessions when the quality points contingency was not available. For this reason, systematic replications of this procedure are needed.

Second, although we provided students with the opportunity to review answer keys with their interteaching partner, this was not a requirement of the course. Therefore, it is unclear if the immediacy of feedback for cooperative performance made a significant difference only during the quality points contingency. Future research on both the availability of quality points and immediacy of feedback for performance are needed before any conclusive statements are made with respect to this component of interteaching.

Third, quizzes in the present study covered relatively little information (e.g., 10-15 pages of reading) compared to the exam probes that have covered more than one topic in previous studies (Saville & Zinn, 2009). Questions included in the probes for the present study were primarily informational and may not have required higher level comprehension or critical thinking. These types of questions were generally reserved for review exams that were administered four times throughout the semester for students in the present study.

Fourth, the quality points in the present study were provided as extra credit. Boyce and Hinline (2002) recommend that quality points account for 10% of students' overall course grade. The decision to incorporate quality points only as bonus was made based on anecdotal reports from students indicating they experienced less stress and anxiety when final course grades were not adversely impacted by a peer's performance (e.g., when the quality points contingency was not in effect). Future studies should evaluate the relative impact of quality points when the contingency is incorporated as part of students' final grade or as an opportunity for extra credit; and also determine if the value of quality points impacts performance during probes (e.g., by making the overall percentage of quality points higher on each probe).

Fifth, class size in the present study was relatively small. Although some support exists for the effectiveness of interteaching with large class sections (Scoboria & Pascual-Leone, 2009), comparison studies of small vs. large class size have yet to be conducted. Students in smaller class sections typically have the advantage of receiving more individual time from the instructor. Individual attention from the instructor may have impacted student performance in the present study, but this variable was held constant (in effect during both conditions) and should not have impacted the results.

Finally, considering the relatively high distribution of final course grades (9 of 11 students [82%] received a grade of either A or B), it is possible that the decline in average quiz performance for the last quiz of the semester (#17) was due in part to student knowledge of high overall course grades. As a result, the perceived value of individual and overall quiz performance and quality points may have decreased at the end of the semester.

Given the limited data reported to date on the impact of each individual element of interteaching, future studies should continue to conduct component analyses (e.g., pair discussion, record sheets, etc.). Future research should manipulate the value of quality points to determine if there is a difference in performance when these points are offered as extra credit or as part of the overall course grade. The quality points contingency could also be set at a higher level for each test probe, or offered for a single question that involves the use of critical thinking. In addition, future research may enhance the notification of feedback and attainment of quality points in an immediate way through on-line notification or delivery of quizzes on-line with immediate feedback on performance. Finally, direct and systematic replications of these procedures in distance education and on-line courses will help to further provide evidence for the robustness of this teaching paradigm.

## Appendices

### Appendix A. Sample quiz

Lesson 1      Form A

Grade \_\_\_\_\_

Name \_\_\_\_\_

Date \_\_\_\_\_

- 
1. Dr. Schmidt consulted with Jimmy. Jimmy wanted to get over being shy. Dr. Schmidt told Jimmy to look other people in the eye, talk loudly, and smile. When Jimmy did so, Dr. Schmidt praised him. Dr. Schmidt was helping Jimmy by using the behavioral science of \_\_\_\_\_.
  2. The person who founded behavior analysis is \_\_\_\_\_.
  3. The science that studies environmental events that change behavior is called \_\_\_\_\_.
  4. Dr. Schmidt consulted with Jimmy. Jimmy wanted to get over being shy. Dr. Schmidt told Jimmy that no one had ever taught him to be direct with other people. Dr. Schmidt explained Jimmy's shyness by using the principle of \_\_\_\_\_ events.
  5. The problem with explaining behavior with private events is that we still must \_\_\_\_\_ the private events.
  6. The first strategy taught in this book is to define your problem as a behavioral \_\_\_\_\_ problem and to gather information about it. I call this the \_\_\_\_\_ strategy.
  7. Behavior analysts see most behavior as \_\_\_\_\_ behavior.
  8. The principle of public events is to look for the causes of behavior in \_\_\_\_\_ events.
  9. Dr. Schmidt consulted with Jimmy about his shyness. Dr. Schmidt told Jimmy that his shyness was the result of no one teaching him to look others in the eye, talk loud enough, and smile. By defining Jimmy's problem this way, Dr. Schmidt was using the \_\_\_\_\_ strategy.
  10. Applied behavior analysis began widespread publication in \_\_\_\_\_.
  11. Short-answer question A. Define *behavior analysis*:  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

From *Principles of Everyday Behavior Analysis (4th ed.)* by L. Keith Miller, pp. 507. © 2006 Toronto: Thomson Wadsworth. Reprinted with permission.



*Appendix B. Sample Preparation Guide.*

**PREPARATION GUIDE #2**

Pair discussion on: **Monday, January 31<sup>st</sup>**

Based on: **Miller: Lessons 2 & 3**

Based on: **Lesson 2**

1. What is *behavior*? What are the four requirements for explaining behavior? Does *behavior* refer only to events or activities that are observable? Explain your answer, and give some examples of behavior that are unobservable. Which of the following would be examples of behavior: (a) running, (b) thinking, (c) perception, (d) memory, (e) eating, (f) emotion, (g) glancing, (h) knowledge, (i) dating, and (j) personality? What seems to distinguish behavior from “non-behavior”?
2. Do behavior analysts deny that thinking exist? Why or why not? Is thinking a public behavior, a private behavior, or both? What is one way that “thinking” could be observed and measured, in other words considered a behavior?
3. What is a behavioral definition? What is the importance of included and excluded behavior when constructing a behavioral definition? What are the two benefits of clear behavioral definitions? What are some possible explanations for inaccurate behavioral definitions?
4. What are self-reports? What are common forms of self-reports? What are some potential problems with self-reports? Why do behavior analysts avoid self-reports? What type of behavioral data do behavior analysts prefer?
5. What is the principle of direct observation? What are some key differences between self-reports and direct observations? Is memory and recall of behavior necessary with direct observation? Why or why not? Which method of behavioral data collection is the most accurate, self-reports or direct observations?

Based on: **Lesson 3**

6. What are considered “uniform” behaviors? What are “nonuniform” behaviors? What are some examples of both uniform and nonuniform behaviors?
7. What is outcome recording? When is outcome recording useful? What are some examples of behavior that leave a “result” behind? Could outcome recording to be used to measure snow shoveling behavior? Why or why not?
8. What is event recording? When is event recording used? Can event recording be used for simple behavior? Complex behaviors? Why or why not?
9. What is an interval type method of observation? What are some novel examples of nonuniform behaviors in which interval recording could be used? What are continuous intervals when using interval recording?
10. What are the differences between interval and outcome recording? What are the main aspects of interval recording, in particular?
11. What is time sample recording? What type of interval is a time sample recording method most often used and why? What is an example behavior in which time sample recording would be appropriate?

## References

- Alvero, A. M., Bucklin, B. R., & Austin, J. (2001). An objective review of the effectiveness and essential characteristics of performance feedback. *Journal of Organizational Behavior Management, 21*(1), 3-29.
- Arntzen, E., & Hoium, K. (2011). On the effectiveness of interteaching. *The Behavior Analyst Today, 11*, 155-160.
- Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., & Parker, R. Academic benefits of peer tutoring: A meta-analytic review of single-case research. *School Psychology Review, 42*(1), 39-55.
- 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.
- Daniels, A. C., & Bailey, J. S. (2014). *Performance management: Changing behavior that drives organizational effectiveness (5th ed.)*. Atlanta, GA: Performance Management Publications.
- Lattal, K. A. (1993). Delayed reinforcement of operant behavior. *Journal of the Experimental Analysis of Behavior, 1*, 129-139.
- Miller, L. K. (2006). *Principles of everyday behavior analysis (4<sup>th</sup> ed.)*. Toronto: Thomson Wadsworth.
- Renner, K. E. (1964). Delay of reinforcement: A historical review. *Psychological Review, 61*, 341-361.
- Roscoe, E. M., Fisher, W. M., Glover, A. C., & Volkert, V. M. (2006). Evaluating the relative effects of feedback and contingent money for staff training of stimulus preference assessments. *Journal of Applied Behavior Analysis, 39*, 63-77. doi:10.1901/jaba.2006.7-05
- Saville, B. K., Cox, T., O'Brien, S., & Vanderveldt, A. (2011). Interteaching: The impact of lectures on student performance. *Journal of Applied Behavior Analysis, 44*, 937-941. doi:10.1901/jaba.2011.44-937
- Saville, B. K., & Zinn, T. E. (2009). Interteaching: The effects of quality points on exam scores. *Journal of Applied Behavior Analysis, 42*, 369-374. doi:10.1901/jaba.2009.42-369
- 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. doi:10.1901/jaba.2006.42-05

Rosales, R., Soldner, J.L., & Crimando, W.

Scheeler, M. C., McAfee, J. K., Ruhl, K. L., Lee, D. L. (2006). Effects of corrective feedback delivered via wireless technology on preservice teacher performance and student behavior. *Teacher Education and Special Education, 29*(1), 12-25.

Scoboria, A. & Pascual-Leone, A. (2009). An ‘interteaching’ informed approach to instructing large undergraduate classes. *Journal of the Scholarship of Teaching and Learning, 9*, 29-37.

Solomon, B. G., Klein, S. A., & Politylo, B. C. (2012). The effect of performance feedback on teachers’ treatment integrity: A meta-analysis of the single-case literature. *School Psychology Review, 41*(2), 160-175.