

## MIT Open Access Articles

### *An Evaluation of the Pair Discussion Component of Interteaching*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**As Published:** <https://doi.org/10.1007/s40732-018-0269-0>

**Publisher:** Springer International Publishing

**Persistent URL:** <https://hdl.handle.net/1721.1/131469>

**Version:** Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

**Terms of Use:** Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



An Evaluation of the Pair Discussion Component of Interteaching

Rocío Rosales

University of Massachusetts Lowell

and

James L. Soldner

University of Massachusetts Boston

and

Libin Zhang

Massachusetts Institute of Technology

**Author Note**

We thank Mathew Swerdan, Joseph Veneziano, and Beatriz Querol Isabelle Del Rosario for assistance with data analysis, and Anya Weber for her editorial assistance with this manuscript.

Address correspondence to either Rocío Rosales: 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) or James L. Soldner: School for Global Inclusion and Social

Development, University of Massachusetts Boston, 100 Morrissey Boulevard, Boston, MA

02125, [james.soldner@umb.edu](mailto:james.soldner@umb.edu)

An Evaluation of the Pair Discussion Component of Interteaching

**Abstract**

The purpose of the present study was to evaluate the impact of the pair discussion component of interteaching on student quiz performance for two sections of an introductory undergraduate course in behavior analysis with a total of 49 students. An alternating treatments design was used whereby the pair discussion was alternated in a quasi-random fashion with a whole-class discussion throughout the semester. In both experimental conditions, all other components of interteaching were in effect. Feedback on quiz performance was provided immediately upon quiz submission. Results show a slight advantage for the pair discussion condition, but no statistically significant differences between the two conditions. Social validity results show a preference for whole-class discussion. These results are discussed in light of the limitations and strengths of the study, and we outline directions for future research.

*Keywords:* interteaching, component analysis, pair discussion, college teaching

Traditional classroom lecture has been reported as the predominant method of college instruction (Benjamin, 2002), despite its relative ineffectiveness in promoting student learning and satisfaction (McKeachie & Svinicki, 2006). Interteaching (Boyce & Hineline, 2002) offers an effective alternative to lecture (Soldner, Rosales, & Crimando, 2015). This behavioral teaching method has been repeatedly demonstrated to enhance student success and satisfaction in the classroom (see Querol, Rosales, & Soldner, 2015 for a comprehensive review of empirical studies on interteaching). For example, Saville, Zinn, Neef, Van Norman and Ferreri (2006) were the first to systematically examine the effectiveness of interteaching relative to classroom lecture through two experimental studies. The first study was comprised of students enrolled in a graduate-level special education course. The students were administered quizzes following alternating conditions of interteaching and lecture. Results demonstrated that quiz scores were higher following interteaching when compared to lecture. The second study included students in two sections of an undergraduate research methods course, with the two teaching conditions alternated and counterbalanced across sections of the course. Results of this study also demonstrated superior performance on quiz scores following interteaching relative to lecture. In addition, both studies reported that the majority of students preferred interteaching.

Interteaching is comprised of several key components. These include: 1) **preparation guides** corresponding to assigned reading material to be completed outside of class; 2) **pair discussions** that take place during class time to review the completed prep guides with a classmate, during which time the instructor serves as facilitator to the discussions; 3) **record sheets** to provide feedback to the instructor on the quality of the pair discussion and to identify challenging class topics; 4) **clarifying lectures** that follow each respective pair discussion and are used to review the most challenging class topics listed on the record sheets; 5) frequent **test probes** based directly on the prep guide material to evaluate student learning; and 6) **quality points**, an explicit cooperative contingency whereby additional

## Running head: PAIR DISCUSSION

points are delivered to students on a test probe if both students in a pair meet a predetermined criterion on any given question(s) or an entire probe (Boyce & Hinline, 2002).

Despite some promising findings, the literature on the relative impact of each component of interteaching is limited. Existing studies have evaluated the relative impact of quality points, clarifying lectures, prep guides, and the pair discussion components of interteaching, with these studies producing largely different results. For example, Saville and Zinn (2009) evaluated the impact of the quality points component on exam scores for students in two sections of an introductory psychology course. Results showed the presence of quality points did not significantly impact student performance on the exams when two conditions were alternated: interteaching with quality points, and interteaching with no quality points. In contrast, Rosales, Soldner, and Crimando (2014) showed quality points made a difference if feedback was delivered more immediately to students. In this study, the instructor provided answers to in-class quizzes immediately upon submission of the quiz.

Saville, Cox, O'Brien, and Vanderveldt (2011) found that the clarifying lecture improved performance, but the differences were not significant. The prep guide component was investigated by Cannella-Malone, Axe, and Parker (2009) by comparing student performance on quizzes that were preceded by either a prep guide with questions developed by the students, or a prep guide with questions developed by the instructor of the course. Results showed participants scored slightly higher on quizzes preceded by student-generated prep guides. However, there were differences based on the types of questions, with answers to problem-solving short-answer questions consistently better when students were required to generate their own prep guide questions. Collectively, these studies have evaluated three of the six components described by the seminal article on interteaching (Boyce & Hinline, 2002).

Boyce and Hinline (2002) defined interteaching as a "mutually probing, mutually informing conversation between two people" (p. 220). This definition emphasizes the importance of the pair

## Running head: PAIR DISCUSSION

discussion. Grouping students in pairs when interteaching is implemented in the learning environment may help to control for the possibility of students underpreparing for in-class discussions with a classmate. For example, there is evidence that when students are placed into large groups there is a higher likelihood of “social loafing” (Chidambarum & Tung, 2005; Latane, 1981). From a behavioral perspective, social loafing can be interpreted as an independent group contingency. That is, when reinforcement is available to all members of a group regardless of individual performance by all group members, the response effort of each individual group member may decrease (Litoe & Pumroy, 1975). Alternatively, an interdependent group contingency requires each individual member of a group to perform at predetermined levels in order to gain access to reinforcement. The pair discussion, and specifically the availability of quality points in interteaching, can be described as an interdependent group contingency. Hence, the possibility of social loafing may be remedied when students are placed into pairs for discussion. Furthermore, there is enhanced individual accountability and a “social” contingency when students discuss in pairs, such that students can more readily gauge partner preparedness to discuss the material during class time.

However, there are also challenges in placing students into pairs to discuss material. Instructors may be hesitant to incorporate this component because the time spent with each student pair may be unequal or inadequate, especially in sections with higher student enrollment (Scoboria & Pascual-Leone, 2009). Other researchers have reported strong preferences among students for working in larger groups (possibly due to students’ familiarity with a group-based classroom lecture format), thereby excluding the pair discussion component (Goto & Schneider, 2010).

To our knowledge, only one study to date has directly evaluated the impact of the pair discussion component in interteaching, while another looked at the impact of group size during the pair discussion. Truelove, Saville, and Patten (2013) evaluated the effectiveness of student performance in interteaching when group size during pair discussion was manipulated (groups of two vs. four). The

## Running head: PAIR DISCUSSION

researchers reported no statistically significant differences between the students' quiz scores on the sections of the six exams administered, a cumulative final, or the total number of points earned in the semester. However, because a no-discussion condition was excluded from the study, the possibility that the pair discussion component of interteaching is not relevant to its effectiveness cannot be ruled out.

Most recently, Soldner, Rosales, Crimando, and Schultz (in press) conducted a component analysis of the pair discussion component on interteaching in an online graduate rehabilitation course. In this study, two conditions were implemented with the same group of students in a quasi-random fashion throughout the semester. The first condition included all key components of interteaching, including the pair discussion. For class sessions in which pair discussion occurred, prep guides were made available to students at least one week before the class session when students were expected to discuss the material with a classmate. During each class session, students were each assigned by the instructor to work with one classmate in an online breakout room via the web platform *Adobe Connect* for pair discussion. During discussion time, the instructor periodically entered individual breakout rooms to facilitate the pair discussion, answer questions, and ensure the students stayed on topic. Following each discussion, students individually completed an online quiz based on the prep guide material. Students received participation points for engaging in pair discussion and submitting a record sheet electronically at the end of each discussion. One quality point was made available to student pairs when both participants achieved a score of 90% or higher on their respective quizzes for that class. The second condition was identical to the first, except that there was no pair discussion or quality points contingency. During class time, students were instructed to review the prep guide individually in an online breakout room. The instructor periodically entered individual breakout rooms to answer questions on the assigned class topics. Students submitted a record sheet at the end of each class session, and received participation points for attending and for studying the prep guide on their own during class time.



Results of this study showed the pair discussion condition resulted in higher quiz scores ( $p < .01$ ). Individual student performance from each quiz indicated that 22 out of 25 students performed better when a pair discussion was incorporated into the class. Additionally, social validity findings from student reports collected at the end of the semester indicated that the majority of students preferred interteaching with the pair discussion, and also that they felt that they learned more during interteaching sessions with the pair discussion condition than without pair discussions. This component analysis provides initial empirical evidence for the inclusion of pair discussion as a part of interteaching, but further study is needed.

Given the paucity of research on the pair discussion component, the purpose of the present study was to systematically evaluate the impact of the pair discussion on student performance and satisfaction when interteaching was used in the traditional, face-to-face college classroom. Participants were students in two course sections of an undergraduate course in behavior analysis taught by the same instructor (the first author). The dependent measure was student quiz performance on chapter quizzes administered following the discussion component. Class sessions were assigned to one of two conditions: interteaching with pair discussions, and interteaching with whole-class discussions. All other components of interteaching were in effect during both experimental conditions (prep guides, clarifying lectures, and quality points).

## **Method**

### **Participants and Setting**

Forty-nine students from two sections of an undergraduate course in applied behavior analysis provided informed consent to serve as participants in this study. The classes met twice a week for 75 minutes during each class session. There were 25 students in Section 1 (5 males and 20 females). The age of students in this section ranged from 20 to 27 ( $M = 21.66$ ,  $SD = 1.54$ ). The class was comprised of upperclassmen (10 juniors and 15 seniors), with a self-reported GPA ranging from 2.3 to 3.9 ( $M = 3.25$ ,

## Running head: PAIR DISCUSSION

$SD = 0.34$ , see Table 1). There were 24 students in Section 2 (11 males and 13 females). The age of students in this section ranged from 20 to 25 ( $M = 21.38$ ,  $SD = 1.20$ ). The class was comprised of upperclassmen (8 juniors and 16 seniors), with a self-reported GPA ranging from 2.5 to 3.8 ( $M = 3.37$ ,  $SD = 0.36$ , see Table 1). We conducted descriptive analyses on each of these variables, reported means and standard deviations for continuous variables, and numbers for categorical variables. Other than gender distributions for each of the class sections, the groups were not significantly different (all  $ps > .05$ ) on any of the measures.

### **Experimental Design, Dependent Variable, and Interobserver Agreement**

An alternating treatments design was used in which pair discussion was alternated with a whole-class discussion. The two conditions were counterbalanced across the two sections of the course and presented in a quasi-random fashion throughout the semester, with the constraint that each condition could occur for no more than three consecutive class sessions. That is, if a class session in the first section was assigned to the pair discussion condition, the second section was assigned to the whole-class discussion condition on the same day. The objective was to ensure that students in either section had an equal number of opportunities to experience the two conditions, and that each prep guide was assigned to both conditions, albeit in two separate sections of the same course.

A total of 10 quizzes were administered per class section, and five were assigned to pair discussion, while the other five were assigned to the whole-discussion condition in an alternating format. Students were notified of the condition in place at the beginning of each class session. This was done to prevent any potential a priori differences in students' preparation before arriving to class. That is, if students were privy to the schedule of assigned sessions at the beginning of the semester, it is possible they might have prepared more for classes that were designated as "pair discussion" sessions than for classes designated as "whole-class" discussion sessions. Class attendance was recorded with the submission of a record sheet at the end of each class period. These records did not indicate a

## Running head: PAIR DISCUSSION

difference in the number of students present for each condition at the end of the semester. In addition, the possibility that students would only attend classes assigned to one of two conditions was partially addressed by informing students of the condition in place only after class began.

The primary dependent measure was average performance on 12-point quizzes administered immediately following each condition. A total of 10 quizzes were administered throughout the semester in each section of the course. Quiz scores included in this analysis comprised 19% of the students' final grade in the course. The first author developed all the quizzes based directly on the information from the prep guides. That is, there was high correspondence between the prep guide questions and quiz questions for each assigned prep guide. The quizzes included multiple-choice, true-false, fill-in-the-blank, and short answer questions. Each quiz was comprised of the same number of question types to equate the level of difficulty across quizzes.

A secondary measure consisted of a social validity questionnaire that included two questions: 1) Overall, which method did you like better? and 2) Overall, which method did you feel you learned the most with? Students were asked to respond to each question by selecting one of the following three options: A) pair discussion (working with one other student), B) whole-class discussion, or C) no preference.

Two independent graders scored 16 of the 20 quizzes (8 per class; 80% of total quizzes administered). Before the quizzes were graded, the student's name was omitted and a copy was made to prevent any potential instructor bias. The instructor of the course served as the primary grader for each quiz, and a secondary grader (a graduate teaching assistant for the course) was provided with an answer key developed by the instructor to score all quizzes. Interobserver agreement (IOA) was calculated by summing the number of agreements and dividing by the number of agreements plus disagreements and multiplying by 100. Mean overall IOA was 88.5% (range 70–100%). It should be noted that the low rates of disagreement were exclusively based on the short-answer responses

## Running head: PAIR DISCUSSION

provided by students on the quizzes designed for this course. When the secondary grader scored less than 80% IOA on any given quiz, the primary grader reviewed the response options for the short-answer questions and IOA improved following this additional training.

### Procedure

The procedure for this study was as follows: preparation guides and guided notes were made available to students on Blackboard, the learning management system used for the class, at least one week before each due date. Students were responsible for completing the prep guide before coming to class. At the start of class, the instructor delivered a clarifying lecture on the previous class session's topic. This lecture lasted approximately 25 minutes. The remainder of the class time was devoted to a 20- to 25-minute discussion of the prep guide due for that class session, completion of the record sheet, and completion of an assigned 12-point quiz.

All students present in class each day were randomly assigned to one of two conditions depending on the class day: during the pair discussion condition, all students were assigned to work with one other classmate; during the whole-class discussion condition, students were invited and encouraged to ask and answer questions related to the prep guide during a discussion that was facilitated by the instructor. Additional details on each condition are presented below.

**Pair discussion.** For class sessions assigned to the pair discussion condition, students were placed into dyads by the instructor. Assignment of pairs was made in a quasi-random fashion by selecting students whose first or last names began with the same letter of the alphabet, by the color of clothing worn that day, or seating arrangement in the class (e.g., someone sitting to the left or right, someone sitting across the room), and students were occasionally permitted to select their own partner for the day. Students were instructed to review the prep guide with their partner in detail and to discuss all of the questions even if they agreed on a response. Students were given 20–25 minutes to complete these discussions. Following each discussion, students completed a record sheet to provide feedback to the instructor on the quality of the discussion and ask for clarification or further help on topics they found most challenging. The record sheets were submitted to the instructor at the end of

## Running head: PAIR DISCUSSION

each class period, and students were awarded participation points for each completed record sheet. Participation points comprised 25% of the students' final grade.

The information gathered from the record sheets was used to create a clarifying lecture for the next class session, as described above. Following completion of each record sheet, students were administered a 12-point in-class quiz. Quizzes covered the information from the prep guide discussed during each class session. Quality points were awarded to both students if they both scored 83% (10/12 points) or higher on a quiz. This explicit, interdependent contingency resulted in two bonus points added to each student's quiz score. If either student received a score below this criterion, neither was awarded quality points. Individual student bonus points were compiled as a separate point total from individual and total quiz score totals, and were added to the student's overall course grade at the end of the semester. The total possible quality points accounted for an additional 6% of the course grade. The instructor provided immediate feedback on student performance following every quiz via delivery of an answer key upon quiz submission. We opted to provide immediate feedback in this manner as a way to enhance the potential value and effectiveness of the quality points (Rosales et al., 2014).

**Whole-class discussion.** All of the same procedures outlined in the pair discussion condition above were implemented during the whole-class discussion condition, with the exception that the instructor did not place students in dyads, but instead facilitated discussion on the prep guide with the entire class. Students were still expected to complete the prep guide before the start of class, and also completed a record sheet before the end of class. This record sheet was submitted to the instructor, and participation points were assigned in the same manner described for the pair discussion condition. The whole-class discussion also lasted 20–25 minutes. The instructor started the discussion by asking students if they had any questions on the prep guide. If a student asked a question that was drawn directly from the prep guide for that class period, the instructor posed the question to the class and invited responses from any student present. Although we did not collect data on class participation

during these sessions, we can report anecdotally that a small number of students participated in the whole-class discussions. In this regard, the whole-class discussion was similar to, but not identical, to a lecture. That is, the instructor stood at the front of the classroom and facilitated discussion amongst the class as a whole, but this discussion was focused on the prep guide that had been distributed a week before the scheduled class session. Therefore, some but not all of the components of interteaching were still in effect during these class sessions. Students were also encouraged to ask follow-up questions, and the instructor selected a few questions to review in detail with the entire class, but the prep guide was not reviewed from start to finish unless specifically requested from students. This never occurred. Quizzes were administered in the same manner as described above, with the exception that bonus points (instead of quality points) were made available to students if they individually scored at least an 83% on the assigned quiz. That is, an independent contingency was in place during this condition.

### Results

A total of 10 quizzes were administered in each section of the class. The results of average quiz performance across the two class sections for the alternating conditions showed that overall, students performed slightly better throughout the semester on quizzes preceded by pair discussion ( $M = 9.14$ ,  $SD = 1.21$ ) compared to quizzes preceded by whole-class discussions ( $M = 8.99$ ,  $SD = 1.91$ ). However, results of a paired two sample for means t-test revealed that the difference in quiz scores was not statistically significant ( $t_9 = 1.116$ ,  $p = .293$ ).

The mean score difference for individual performance is presented in Figure 1. These results indicate that 12 out of 25 students' average performance was better on quizzes that were preceded by the pair discussions in Section 1, and 10 of 24 students' average performance was better on quizzes preceded by this condition in Section 2. We also evaluated the difference in individual average performance between the two conditions. The difference was not statistically significant:  $t_{25} = 0.27$ ,  $p =$

0.78 (Section 1) and  $t_{24} = -0.50, p = 0.61$  (Section 2). Taken together, the Mann-Whitney U test shows that the individual average performance between the pair-discussion and whole-class discussion was not significantly different ( $U = 46, p = .76$ ). The results of the social validity measure showed that students in both class sections reported preference for working as a whole class and perceived better learning outcomes following this condition when compared to the pair discussion (see Table 2).

The social validity questionnaire used in this study included a section where students could write in responses to explain their selection. This qualitative data showed that many participants reported that their peers were not as knowledgeable as the instructor, which made them less trusting of the information they generated together. In addition, some participants reported that they preferred the whole-class discussion because often their assigned partner was underprepared for the interteach discussion, which in turn left them feeling underprepared for the quiz that was completed at the end of each class session.

### **Discussion**

Previous research has shown an overall advantage for interteaching over traditional college instruction when it comes to student performance (Saville, Zinn, & Elliot, 2005; Saville et al., 2006). Some studies have conducted component analyses to determine the relative impact of the components of interteaching, including the clarifying lecture (Saville et al., 2011), quality points (Rosales et al., 2014; Saville & Zinn, 2009), and pair discussion (Soldner et al., in press). This is the first study to empirically evaluate the impact of the pair discussion component of interteaching in a face-to-face undergraduate course. Results from the present study revealed low average performance for all participants, and the difference between the two conditions was not statistically significant. These results suggest that although the discussion component of interteaching may be an important feature of this instructional technique, discussion may not need to occur in pairs. In addition, the social validity measures indicate that students preferred the whole-class discussions.



Several factors should be considered in the interpretation of these results. First, a contrast effect may help account for the differences reported. That is, students experienced the two discussion conditions in a rapidly alternating format, and one week of class may have included one class session with pair discussion and one class session with whole-class discussion. Future studies on this topic should include a no-discussion control condition similar to that described by Soldner et al. (in press). This set-up would not preclude students from asking individual questions of instructors during class, but it would also not be required of them to engage in small-group or pair-discussions at any time. There are possible methodological and ethical concerns associated with this type of control condition in a traditional classroom environment. That is, students paying tuition and fees for a face-to-face class may perceive a control condition of this kind as an inadequate learning environment. It may also be difficult for instructors to speak to students on an individual basis in a single class period. An alternative is to conduct a component analysis of the pair discussion component of interteaching in a laboratory setting (e.g., Saville et al., 2005).

A second limitation of this study is that the majority of participants were upperclassmen (i.e., classified as juniors and seniors). This student demographic typically has a longer history with lecture-based courses. The unfamiliarity with the course structure may help account for the reported *preference* for whole-class discussions at the end of the semester, because this format resembles a lecture-based course where instructors attempt interactive discussions that are still limited to participation by only a handful of students. Anecdotally, we saw a similar pattern of interaction emerge during the whole-class discussion sessions.

Third, the primary dependent variable (quiz scores) comprised a relatively small amount of possible points in the class over the course of the semester. It is possible that the low value of this reinforcer had an impact on student preparedness and performance on quizzes. Future studies should

## Running head: PAIR DISCUSSION

consider manipulating the value of quiz scores if these serve as the primary dependent variable. To date, empirical studies on interteaching have largely focused on quiz or exam performance as the primary dependent measure. College courses often include other ways to evaluate student performance and learning. Instructors and students alike may be most concerned with students' ability to apply classroom knowledge in practice, especially for courses in human service-related fields. Future research in this area should be designed to demonstrate a difference in performance on specific tasks following a class session with interteaching compared to class sessions with lecture.

Fourth, we did not conduct pre-assessments of student knowledge at the beginning of the semester. Future studies may wish to include this sort of pre-test assessment as has been done in previous studies on this topic (Saville et al., 2006). For example, students could be assigned a pre-test to assess all knowledge to be covered over the course of the semester and administer the same evaluation at the end of the semester. The difference in scores for each individual participant can then be reported. Future studies may also wish to increase the value of quizzes by adding more questions and/or increasing the point value attached to each individual question.

Fifth, since the same instructor taught both sections of the course, the generalizability of these results is limited because these may be dependent at least partially on the instructor's teaching style, strengths, and/or weaknesses. Holding this variable constant is beneficial for the internal validity of the study, but it limits the generalizability of the results. We also did not evaluate long-term retention of the material learned (Saville, Bureau, & Zombakis, 2014). In the present study, we conducted post-discussion quizzes immediately following either the pair discussion or whole-class discussion. The quizzes were based on the material discussed in class the day the quiz was administered. Future studies should include a maintenance probe in the form of cumulative exams that include only the information from either condition (i.e., content covered in the pair discussion condition vs. content covered in the whole-class discussion condition) to determine if there is a difference in retention rates.

Finally, the implementation of interteaching in this study varied from the description conceptualized by Boyce and Hinline (2002). A recent review of interteaching research by Sturme, Dalen, and Fienup (2015) shows that no published studies to date have reported data on treatment integrity. In this study, the variables that differed from the original conceptualization include the percentage of quality points available to students (i.e., 6% of the students' overall course grade), and administration of post-discussion quizzes immediately following the discussion of the prep guide material. Previous research has suggested that quality points should consist of roughly 10% (or even 15–20%) of students' overall course grade (Boyce & Hinline, 2002; Querol et al., 2015). The percentage of available points was held constant across conditions, and the possibility of this confound in this study is therefore minimized. However, future studies should evaluate the value of quality points as a factor in the impact of interteaching in the classroom.

The formation of student groups for each pair discussion was also implemented as suggested by Boyce and Hinline (2002), but given the low number of students per class, there were inevitably times when students worked with the same partner. We aimed for variability in the assignment of discussions throughout the semester, as described above, but did not collect specific data on partner assignments. As Sturme and colleagues (2015) pointed out, the lack of treatment or procedural integrity data reported in previous studies presents difficulties for direct and systematic replications and implementation of interteaching by other instructors. Perhaps most importantly, the differences may impact the outcomes reported by different groups of researchers. Future studies should report specific details on these variables, as this may shed light on subtle but systematic differences between experimental conditions. In addition, a review of the studies that have been published to date that evaluates these differences may be informative for future research.

Other avenues for future research include systematic evaluation of the prep guide component of interteaching. To our knowledge, no published studies have yet been conducted on this topic. Given

that prior component analyses have reported mixed results (Rosales et al., 2014; Saville et al., 2011; Saville & Zinn, 2009; Soldner et al., in press), it seems worthwhile to evaluate the impact of the prep guide. Furthermore, considering the potential variability with the development of prep guides, including the content, format, length, etc., the treatment and/or procedural integrity of this component of interteaching must also be considered and evaluated, as discussed by Sturmev and colleagues (2015).

Finally, future studies should continue to evaluate students' demographic variables and their relationship to performance and preference for instruction that follows an interteaching format. In these studies, we collected certain demographic information, but did not specifically control for demographic variables as a way to evaluate the effectiveness of interteaching with and without the pair discussion component. A previous study by Saville, Pope, Truelove, and Williams (2012) investigated the relations between student GPA performance and demographic variables on exam performance during interteaching and lecture, and found that interteaching improved performance most notably for students with low and moderate GPAs. The participants in the present study had an average GPA over 3.0 and were enrolled at a four-year university. The large majority were upperclassmen and considered "traditional" (ages 18–25) college students. Future studies might evaluate the effectiveness of interteaching and its components with diverse student bodies, and report more specifically on particular strengths of the participants included in the analyses (e.g., GPAs, prior experience with interteaching format).

### **Compliance with Ethical Standards**

**Conflict of Interest:** The authors declare that they have no conflict of interest.

**Ethical approval:** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Informed consent:** Informed consent was obtained from all individual participants included in the study.

### References

- Benjamin, L. T. (2002). Lecturing. In S. Davis & W. Buskist (Eds.), *The teaching of psychology: Essays in honor of Wilbert J. McKeachie and Charles L. Brewer* (pp. 57–67). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Boyce, T. E., & Himeline, P. N. (2002). Interteaching: A strategy for enhancing the user friendliness of behavioral arrangements in the college classroom. *The Behavior Analyst, 25*, 215–226.
- Cannella-Malone, H. I., Axe, J. B., & Parker, E. D. (2009). Interteach preparation: A comparison of the effects of answering versus generating study guide questions on quiz scores. *Journal of the Scholarship of Teaching and Learning, 9*(2), 22–35.
- Chidambarum, L., & Tung, L. L. (2005). Is out of sight, out of mind? An empirical study on social loafing in technology-supported groups. *Information Systems Research, 16*(2), 149–168.  
doi:10.1287/isre.l050.0051
- Goto, K., & Schneider, J. (2010). Learning through teaching: Challenges and opportunities in facilitating student learning in food science and nutrition by using the interteaching approach. *Journal of Food Science Education, 9*, 31–35. doi:10.1111/j.1541-4329.2009.00087.x
- Latane, B. (1981). The psychology of social impact. *American Psychologist, 36*(4), 343–356.
- Litoe, L., & Pumroy, D. K. (1975). A brief review of classroom oriented contingencies. *Journal of Applied Behavior Analysis, 8*(3), 341–347. doi:10.1901/jaba.1975.8-341

Running head: PAIR DISCUSSION

McKeachie, W.J., & Svinicki, M. (2006). *McKeachie's Teaching Tips: Strategies, Research, and Theory for College and University Teachers*. Boston: Houghton Mifflin.

Querol, B. I. D., Rosales, R., & Soldner, J. L. (2015). A comprehensive review of interteaching and its impact on student learning and satisfaction. *Scholarship of Teaching and Learning in Psychology, 1*(4), 390–411. doi:10.1037/stl0000048

Rosales, R., Soldner, J. L., & Crimando, W. (2014). Enhancing the impact of quality points in interteaching. *Journal of the Scholarship of Teaching and Learning, 14*(5), 1–11. doi: <https://doi.org/10.14434/josotly14i5.12746>

Saville, B. K., Bureau, A., Zombakis, J. (2014). Interteaching and lecture: A comparison of long-term recognition memory. *Teaching of Psychology, 41*(4), 325–329. doi: 10.1177/0098628314549704

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., Pope, D., Truelove, J., & Williams, J. (2012). The relation between GPA and exam performance during interteaching and lecture. *The Behavior Analyst Today, 13*, 27–31. doi:10.1037/h0100728

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., & Elliot, M. P. (2005). Interteaching versus traditional methods of instruction: A preliminary analysis. *Teaching of Psychology, 32*, 161–163. doi:10.1207/s15328023top3203\_6

Running head: PAIR DISCUSSION

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

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.

Soldner, J. L., Rosales, R., & Crimando, W. (2015). A comparison of interteaching and classroom lecture in rehabilitation education. *Rehabilitation Counselors and Educators Journal, 8*(1), 91–100.

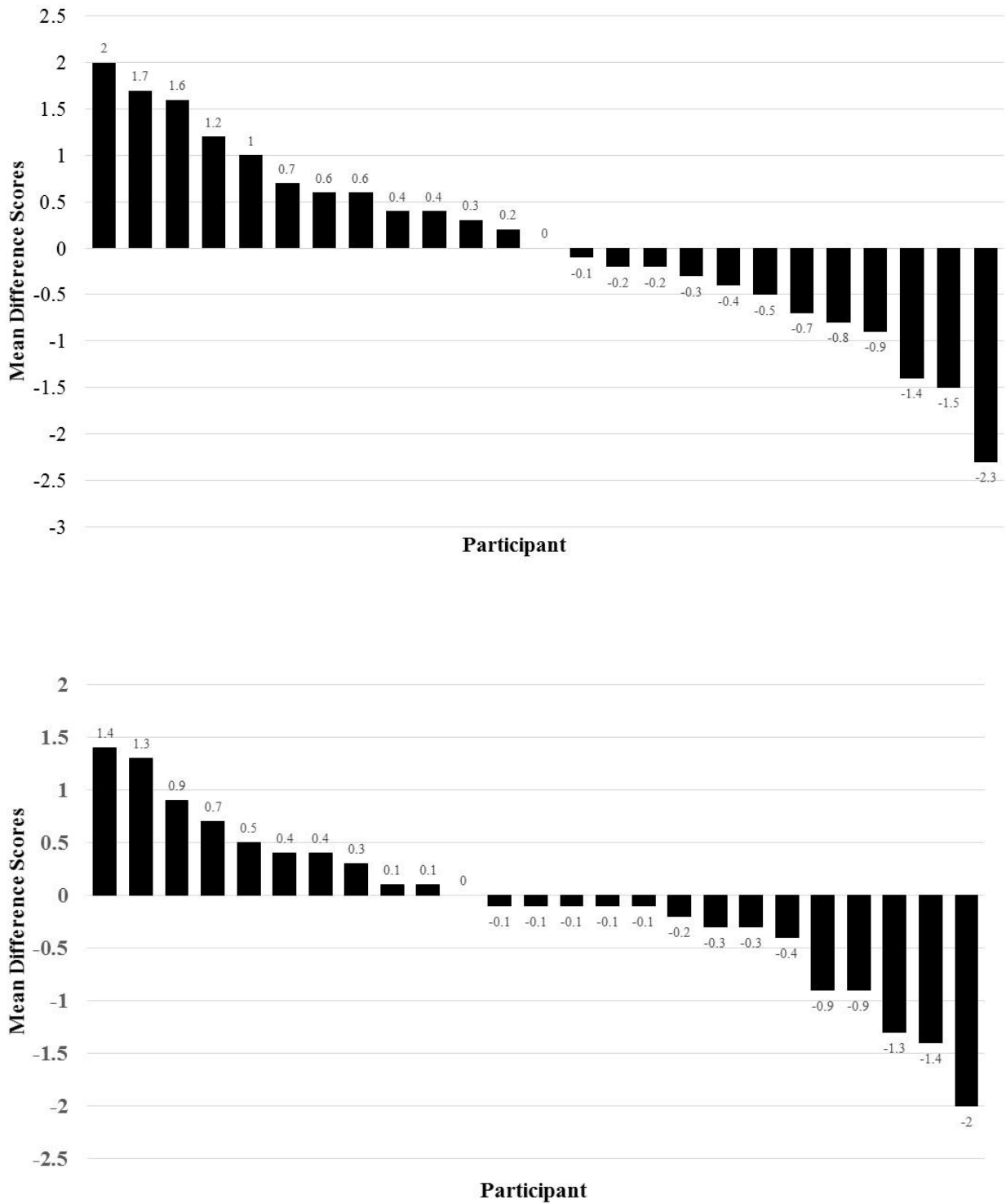
Soldner, J. L., Rosales, R., Crimando, W., & Schultz, J. C. (in press). Interteaching: Application of an empirically supported behavioral teaching method in distance rehabilitation education. *Rehabilitation Research, Policy, and Education*.

Sturmey, P., Dalfen, S., & Fienup, D. M. (2015). Inter-teaching: A systematic review. *European Journal of Behavior Analysis, 16*(1), 121–130. doi:10.1080/15021149.2015.1069655

Truelove, J. C., Saville, B. K., & Van Patten, R. (2013). Interteaching: Discussion group size and course performance. *Journal of the Scholarship of Teaching and Learning, 13*, 23–30.



**Figure 1.** Mean individual difference scores between pair discussion and whole-class discussion for all participants in Section 1 (top panel) and Section 2 (bottom panel).





**Table 1.** Demographic information for students.<sup>1</sup>

	<b>Section 1</b>	<b>Section 2</b>
<b>Sex</b>		
n	25	24
Male	5 (20%)	11 (45.8%)
Female	20 (80%)	13 (54.2%)
<i>M</i> Age	21.66	21.38
<i>M</i> GPA	3.25	3.37
<b>Year in School</b>		
Junior	10 (40%)	8 (33.3%)
Senior	15 (60%)	16 (66.7%)
<b>Employment Status</b>		
Employed	22 (88%)	20 (83.3%)
Not Employed	3 (12%)	4 (16.7%)
<b>Ethnicity</b>		
White	21 (84%)	19 (79.2%)
Hispanic/Latino	2 (8%)	2 (8.3%)
Black/African-American	1 (4%)	1 (4.2%)
Asian	1 (4%)	1 (4.2%)
Native American/Alaskan	0	1 (4.2%)

<sup>1</sup> Although we conducted statistical analyses on each of these measures, we do not report them here. Other than gender distributions for each of the class sections, the groups were not significantly different (all *ps* > .05) on any of the measures.

**Table 2.** Social validity data results for Section 1 (top) and Section 2 (bottom).

Section 1 (N=25)

Question		Percentage (#) of students
<b>1.</b>	Overall, which method did you like better?	
<b>a.</b>	Pair discussions (working with just one other student)	28% (7)
<b>b.</b>	Whole-class discussions	52% (13)
<b>c.</b>	No preference	20% (5)
<b>2.</b>	Overall, which method did you feel you learned the most with?	
<b>a.</b>	Pair discussion (working with just one other student)	20% (5)
<b>b.</b>	Whole-class discussions	60% (15)
<b>c.</b>	No preference	20% (5)

Section 2 (N=24)

Question		Percentage (#) of students
<b>1.</b>	Overall, which method did you like better?	
<b>a.</b>	Pair discussions (working with just one other student)	17% (4)
<b>b.</b>	Whole-class discussions	66% (16)
<b>c.</b>	No preference	17% (4)
<b>2.</b>	Overall, which method did you feel you learned the most with?	
<b>a.</b>	Pair discussion (working with just one other student)	17% (4)
<b>b.</b>	Whole-class discussions	58% (14)
<b>c.</b>	No preference	25% (6)