Interteaching: The Effects of Discussion Group Size on Undergraduate Student Performance and Preference

Michael Gutierrez

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Ilene Schwartz, Chair

Ali Mokdad

Scott Spaulding

Elizabeth West

Program Authorized to Offer Degree:

College of Education

University of Washington

Abstract

Interteaching: The Effects of Discussion Group Size on Undergraduate Student Performance and Preference

Michael Gutierrez

Chair of the Supervisory Committee:
Ilene Schwartz, PhD
Special Education

Interteaching is a college teaching method grounded in the principles of applied behavior analysis. Research on interteaching demonstrates that it improves academic performance, and students report greater satisfaction with interteaching as compared to traditional teaching styles. The current study investigates whether discussion group size, a major component of the interteaching process, affects student performance and satisfaction. On different days, researchers exposed participants to one of three treatment conditions (pairs, small groups, and large groups) and then assessed the effects on performance and satisfaction. The study found no apparent differences on quiz performance based on discussion group size, although students performed better following pair and small group sessions. Students overwhelmingly preferred small groups compared to pairs or large groups. These results are especially relevant for instructors teaching a course with high student enrollment, in both traditional and online courses, where managing groups of two may prove difficult.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Eight years ago, I made a decision to embark on a journey that, I thought, was about obtaining a doctorate degree and exceling in a field that I was more passionate about than anything else in my life. But, it turned into much more. It shaped my future life, introduced me to the family and friends that I will have for a lifetime, and made me into the person that I am proud to be. I am so incredibly grateful for the experience that I have had and could not have accomplished this without the help I have received from so many people in my life.

First and foremost, I would like to thank my wife- for her support, her encouragement, and her willingness to take on so much more than I could ever ask. This selflessness is something that I will always be grateful for. She inspires me every day in so many ways and without her, I wouldn't be where I am today.

I would also like to thank all the parents in my life- Ron, Kathleen, Steve, Holly, Scott, and Hilary. Their phone calls cheering me on, their tokens of advice, and their assurance that my hard work would soon pay off has meant more than they may ever know.

My adviser, Ilene Schwartz, has also been such an important figure over the last seven years. Her patience and mentorship along with an unrivaled knowledge of special education and the importance of inclusive education for all has shaped my character and morals as I continue to develop as a scholar and practitioner in behavior analysis.

Lastly, I would like to thank my friends- my village. They have been with me every step of this process and cheered my highs and supported my lows. Their positivity and their friendship has helped create a path for something that I could only dream of bringing into reality.

DEDICATION

To my wife and stepson,
Courtney and Indiana

Interteaching: The Effects of Discussion Group Size on Student Performance and Preference

Due to the economic downturn of the new millennium, large class sizes have become the new norm within modern higher education institutions (Cullen, 2011). In today's climate, it is not enough that instructors in higher education be well-versed in their content area, the newest instructional strategies, and the latest technologies that can be used to improve learning in the college classroom; they must also be able to adapt these instructional strategies to suit the growing numbers of students in their classes. One empirically-based instructional strategy that has been found to be quite effective in improving academic performance is interteaching, a college teaching strategy grounded in the principles of applied behavior analysis (Boyce & Hineline, 2002). One key component of the interteaching process requires that students be assigned to discussion groups of two while the instructor intermittently checks in with each group as they discuss the assigned material. In Boyce and Hineline's (2002) article introducing interteaching, they state that groups of two are the most optimal because total student engagement and involvement is maintained by natural contingencies of social interaction. They state that working in pairs leave potentially under-performing students no place to hide, requiring every student to participate fully in the discussion. However, they provide no evidence that this is actually the case. While this may be true, it is important that research is conducted to support this claim. Further, this requirement could pose an issue for instructors of classes with high student enrollment. For example, instructors would have to check in with over 20 groups for a class size of 40 or more students. This would certainly prove difficult from the perspective of appropriate classroom management. As a result, interteaching may not as preferred by instructors with high student enrollment. In order to keep interteaching at the forefront of

effective, research-based teaching strategies in the field of higher education, it is essential that interteaching be explored for possible adaptations that would make it an attractive option for instructors with high course enrollment. This view is reiterated by Brown, Killingsworth, and Alavosius (2014) when they state that, "a systematic replication of interteaching protocols within the context of high enrollment courses would provide a scope of utility for instructors charged with educating a large number of students per semester" (p. 136).

This paper will begin by providing a brief overview of the research-based instructional strategies that are embedded within the interteach process, a description of the essential components of interteaching, and a brief literature review of the effectiveness of interteaching across a variety of domains. Then, this study will explore the effects of manipulating group size on various student outcomes including performance, satisfaction, and preference.

Interteaching

The components of interteaching evolved from multiple behavior analytic strategies for classroom instruction. Personalized system of instruction (PSI), first described by Fred Keller in the *Journal of Applied Behavior Analysis* in 1968, involves small self-paced modularized units of instructions where study guides direct learners through pre-arranged modules (Kulik, Carmichael, & Kulik, 1974). Precision teaching, another empirically-based teaching practice works under the assumption that "The learner knows best" and uses directly observable behaviors as its focus when assessing student performance (Lindsley, 1991). Peer tutoring is a system of instruction in which learners help each other and learn by teaching using a variety of roles and contexts (e.g., Delquadri, Greenwood, Whorton, Carta, & Hall, 1986; Goodlad & Beverley, 1989). Cooperative learning, another evidence-based instructional practice embedded within interteaching, asks students to work in small groups to receive rewards or recognition

based on group performance (Slavin, 1980). Lastly, problem-based learning, students work independently on a specific problem or scenario and then work in groups to share their findings and refine their acquired knowledge (Wood, 2003). Each of these evidence-based instructional strategies are present, in one way or the other, in the components of the interteaching process (Saville, Zinn, Neef, Van Norman, & Ferreri, 2006). Given its strong foundation in these behavior analytic approaches, which are grounded in research, an instructor may find interteaching to be quite useful for teaching students effectively.

The purpose of interteaching is to create a formalized process of guiding and reinforcing individual study behavior (Truelove, Saville, & Van Patten, 2013). This is accomplished by providing students with a guide for their pre-class work, creating opportunities for students to discuss their work with peers, and then seeking clarifications by the instructor. Through this process, interteaching specifically identifies what behaviors students should emit to improve overall course performance and then rearranges the reinforcement contingencies to produce those behaviors (Truelove et al., 2013). The interteaching process consists of five steps:

1. Students come to class having completed a preparation guide covering the assigned readings.
2. Students split into dyads to discuss the assigned material and use the prep guide to facilitate discussion while the instructor intermittently checks in with each group.
3. Students fill out a teaching record to seek clarification on specific topics.
4. The following class, a lecture is presented to the class based on the teaching records.
5. A quiz is administered to assess student comprehension.

Specifically, the interteaching process is as follows: first, the teacher constructs a preparation (prep) guide 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. The teacher then distributes the prep guide to students who 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 (more clarification on this later). After the lecture, students break into groups to discuss the preparation guide. During the pair discussions, the teacher (and teaching assistant, if available) move around the classroom, answering questions, guiding conversation, and providing on-going student support. After students have discussed the prep guide thoroughly, they complete a teaching record which provides the teacher with feedback on how the discussions went and which material is difficult to understand. The record sheet is then used to inform the lecture that occurs during the first 30 minutes of the subsequent class. Because the students are clarifying the specific material that they need assistance with, the clarifying lectures become highly effective in helping students better understand the source material (Saville, Cox, O'Brien, & Vanderveldt, 2011).

One of the defining characteristics of interteaching is the process of prep guide, group work, teaching record, clarification, and assessment. Although interteaching has been used across content- such as, nutritional and food science, special education, general psychology, and sociology (Goto & Schneider, 2010; Saville & Zinn, 2009; Saville, Zinn, & Elliot, 2005; Tsui, 2010); contexts- such as, graduate, undergraduate, and controlled laboratory settings (Rosales, Soldner, & Crimando, 2014; Saville et al 2006; Saville, Zinn & Elliot, 2005); and types of learners- such as, low vs. high grade point average (Saville, Pope, Truelove, & Williams, 2012), the process used remains constant. The following section will describe these studies and others as well as the ways in which the interteaching process has been proven effective.

Effectiveness of Interteaching

Multiple studies have shown the effectiveness of interteaching compared to other modes of instruction. It has been found to be more effective on graduate and undergraduate student performance on quiz scores and a cumulative final than traditional lecture (Arntzen & Hoium, 2009; Saville et al., 2006). Both of these studies alternated between lectures and interteaching conditions to test the effects on student performance. Saville et al. (2005) found similar results when comparing interteaching to traditional methods of instruction such as lecture, reading, and a control group for undergraduate students at a moderately sized public university. Within this study, instructors taught students the positive effects of pets on blood pressure by placing them in one of four treatment groups (interteaching, lecture, reading, and control). They found that participants in the interteaching group performed significantly better on a short, multiple-choice quiz compared to the other treatment groups. It was also effective as an organizational and instructional approach for introductory-level sociology courses (Tsui, 2010). In this case study, the author described interteaching and its effectiveness in producing more focused class discussions as well as more focused lecture. In another case study, the authors described how the interteaching method was employed throughout a food science and nutrition class, by describing the interteach process, the tools used, as well as the way in which they measured perceived learning (Goto & Schneider, 2010). Through these measurements- specifically from quizzes and questionnaires, they found that the interteach process fosters critical thinking and enhances student motivation. Another study found that interteaching has a more positive impact on exam performance for students with low and moderate GPAs compared to traditional lecture (Saville, Pope, Truelove, & Williams, 2012). They found this by comparing undergraduate students with high and low GPAs against one another when alternating between interteach and lecture

conditions and taking measurements on exam performance. Lastly, in an essay describing their experience using the interteach process, Scoboria, Sirois, and Pascual-Leone (2009) found that interteaching provides students with a rich rehearsal of material, and emphasizes student engagement, peer discussion, and student/instructor interaction. While this anecdotal information was provided within the broader context of the effectiveness of interteaching, the findings are still pertinent to this study.

Other studies have examined specific components of interteaching and how they may contribute to its effectiveness. Two studies examined the effectiveness of quality points- a component of interteaching that refers to a cooperative contingency in which part of the student's exam grade depends on how well his or her partner performs on specific exam questions (Rosales, Soldner, & Crimando, 2014; Saville & Zinn, 2009). One study found quality points to have a significant impact (Rosales et al., 2014) while the other found no significant impact (Saville & Zinn, 2009). The prep guide component was examined by having students either answer prep guide questions or create the questions themselves (Cannella-Malone, Axe, & Parker, 2009). This study found that quiz performance was not substantially influenced by the way the prep guide was constructed. However, most students preferred answering study guide questions rather than creating the questions, as they felt better prepared for the quizzes. The length of the lecture component of interteaching was analyzed by exposing students to one of three treatment groups-delayed lecture, immediate lecture, or no lecture (Saville et al., 2011). This study found that students who received lectures had consistently higher exam scores and earned significantly more points during the semester. However, whether or not the lecture was delayed or immediate did not seem to matter.

The acceptability and social validity of interteaching has also been examined across a handful of studies. Saville et al. (2006) conducted two studies that measured responses to questionnaires that ranked student preference for interteaching compared to lecture. They found that the majority of students preferred interteaching, with fewer students reporting that they preferred lecture or had no preference. Further they found that students reported that they felt they had learned more of the content during interteach than lecture. Zayac and Paulk (2014) asked students to anonymously complete a six-item questionnaire to assess student's preferences for the instructional format they used when in interteach groups. Based on these responses, they found that the majority of the students preferred lecture-based classes compared to interteaching. Given these mixed findings, it is important to further investigate the social validity of interteaching in this study, and if there are components of interteaching that affect student preference ratings.

One component of interteaching that has not been examined extensively is the size of the discussion group- specifically, how the number of students in each interteaching group may affect student performance or preference. This is an important issue given the increasing size of college classes (Mitchell, Leachman, & Masterson, 2016), and the potential difficulty of managing multiple interteaching groups both live and on line. Boyce and Hineline (2002) hypothesize that groups of two would be the most efficient compared to groups of 3 or more due to an idea of social loafing where students are prone to exert less effort on a task if they are in a group vs. working alone. When in groups of two, the authors state, "total involvement is maintained by natural contingencies of social interaction that are minimally aversive while leaving no place to hide" (p. 219). While this claim indeed has weight, only one study has addressed this component of the model. Truelove et al. (2013) examined whether group size

affects student performance in an interteaching-based course by placing students in groups of two vs. four in an undergraduate psychology course. The authors found no significant differences between groups across 6 unit exams, a cumulative final, and in the total number of points earned across the semester. The authors suggest that performance may deteriorate with increased group size but additional research is needed to determine if this is true. This is especially important because if these findings are accurate and generalizable, teachers may have more flexibility in how they arrange their classroom interteach groups to suit their growing classrooms. The purpose of this study was to replicate and expand upon the Truelove et al. (2013) study by examining the impact of group size- pairs (2 individuals), small groups (4-6 individuals), and large groups (7-9 students) on student quiz scores, a cumulative final, and student preference/satisfaction. Specifically, this study aimed to understand the effects of interteaching group size on student performance, satisfaction, and preference for undergraduate students taking a course in applied behavior analysis.

Method

Participants

Participants were 19 undergraduate students at the University of Washington taking a five-credit introductory course in applied behavior analysis. This class focused on basic behavioral principles, defining behaviors, measuring behaviors, effective strategies to teach new skills, and the ethics involved in behavior change programs. This course fulfilled the coursework requirement to become a Registered Behavior Technician ^{TM} (Behavior Analysis Certification Board [BACB], 2017), but did not include a practicum or field work component. To fulfill the coursework requirements, the class work focused on the specific components outlined by the BACB (See Appendix A).

The course met twice weekly for 2 hours and 20 minutes per class. Mean age for participants was 20.84 years old and ranged from 18 to 32 years of age. The majority of the students were female (89%). Ten students were Caucasian (53%), seven were Asian (37%), one student was Black/African American, and one student did not report ethnicity. The majority of students were majoring in Early Childhood Family Studies (63%), followed by Psychology (26%), English (.05%), and no answer (.05%). Students were Freshmen (.05%), Sophomores (21%), Juniors (32%), Seniors (32%), and two did not answer (11%). Before the start of the class, the students' average GPA was 3.47 and ranged from 2.39 to 3.99. For an overview of demographic data, see Table 1 below:

Table 1.

Participant Demographics

Gender		Academic Year	
Male	2	Freshman	1
Female	17	Sophomore	4
Age (in years)	M = 20.84 Range = 18-32	Junior	6
Ethnicity		Senior	6
Caucasian	10	No Answer	2
Asian	7	Major	
Black/African American	1	Psychology	5
No Answer	1	Early Childhood Fam Studies	12
GPA	M = 3.47 Range = 2.39-3.99	No Answer	2

Materials

Principles of Behavior (Malott & Shane, 2014) was the primary text for this course. Some readings from *The Data Model for Teaching Preschoolers with Autism* (Schwartz, Ashmun, McBride, Scott, & Sandall, 2017) were used to supplement the text. All of the readings came from these two texts. The content for the prep guides and quizzes was based on these readings. One to two chapters were assigned for homework following each class.

There were two special sets of instructional materials that were necessary for this study-preparation guides and teaching records. The purpose of the preparation guides was to guide the reading and comprehension of individual students before they came to class as well as to guide conversation during their interteaching groups. The purpose of the teaching record was to provide the students with an opportunity to ask questions about the content, request items to be addressed in the lecture in the next class, and assess students' satisfaction with the group discussion.

Prep guides were provided to students at least 2 days prior to each class and were created by the researchers. Each prep guide contained 10-29 items and covered 1-2 chapters from the text (20-50 pages). Questions ranged from simple definition questions to complex problem solving scenarios depending on the source material (See Appendix B) and used questions either taken directly from the text book or developed by the researchers. Prep guides were completed and turned in electronically prior to each class via a learning management system used by the university. Students were given a grade of complete/not complete which counted towards 10% of their final grade. Students brought completed copies of the prep guide with them to class. The content of the prep guide guided the small group discussions.

The purpose of the teaching record was to provide students with the opportunity to ask questions related to the learning material. Further, the teaching records were used to measure the level of satisfaction of their interteaching group. These measures were assessed through various questions about the quality and usefulness of the discussion group. Items within the teaching record included the following: duration of the discussion, if sufficient time was provided (yes/no), level of difficulty of the material, and a list of issues that the student would like clarified in the following class lecture. Further, students were asked provide a rating to measure

their preference for the group. Students submitted teaching records at the end of every class and received 5 points for every teaching record they completed, for a total of 10% of the total course grade (See Appendix C).

Two additional questionnaires were used at the beginning and end of the course. These questionnaires were not part of the inter teaching process, but were related to the experimental analysis of interteaching. The first questionnaire collected demographic data including age, gender, academic year, and current GPA to determine how these variables may affect performance across groups. The second questionnaire collected additional social validity data including preference for each group size, quality of each group size, and the pros and cons for group size. Additional space was allowed for students to provide comments about their experience in each group.

Quizzes were administered to assess student performance. Each quiz had 10 multiple choice items that were directly taken from a sample of the material covered in the preparation guides. Quizzes were administered via the online course management system that was used by the university. Students could complete the quizzes on phones, tablets or computers. The instructor arranged to have other computers available to students who needed them. Students were also able to take the quiz on paper upon request. An end of the year, a final exam was administered and each question was coded to determine under what condition the material was addressed.

Procedure

Prior to each class (2-5 days) a prep guide was assigned for students to complete before the beginning of each class. The prep guides covered 20-50 pages of textbook material and was constructed to guide students through the assigned readings of the class. Depending on the

assigned readings, prep guides contained anywhere from 10-29 items and covered items directly from the reading. Students were required to turn in their completed prep guides prior to the beginning of class and came to class prepared to discuss the content. At the beginning of each class, the students heard a brief clarifying lecture that reviewed items discussed on the teaching records filled out during the previous class and then were quizzed on that material. Then, students were randomly split into pairs, small groups, or large groups, depending on the treatment condition, by having each student count off by the number of groups required for that class. Students then sat with the students whose number they matched with. Each class was randomly assigned to one of three potential conditions: In the first condition (Group A) students split into pairs. In the second condition (Group B) students split into small groups (Groups of 4, 5, or 6). In the third condition (Group C) students split into large groups (Groups of 7, 8, or 9). When groups did not split evenly, left over students were placed into one or more of the groups. So on some occasions, groups of 3, 5 or 9 existed depending on the day.

During each group, students discussed material at their own pace using the prep guide as a visual accommodation. These discussions lasted for approximately 30 minutes. During the discussions, the instructor moved around the classroom, answering questions, and guiding the discussions. Following the discussion, students individually completed a teaching record sheet which contained the following pieces of information: (a) student name; (b) partner(s) names; (c) the date; (d) a rating of how well their discussion went (scale of 1-5); (e) statement(s) as to why they thought the discussion did or did not go well; (f) statement(s) on what specific topics or content they would like increased clarification.

Students who participated in the discussions and completed a record sheet earned a small number of participation points across the quarter which totaled 10% of their course grade. At the

beginning of the following class, the instructor lectured over the material that students had listed on the record sheets. Following this clarification lecture, a quiz was administered that covered that material. Then, students returned to pairs, small groups, or large groups (depending on the class that has been randomly assigned) to discuss the assigned prep guide for that class. At the end of the quarter, students took a cumulative final exam that contained items from each of the prep guides to examine how group size may affect knowledge of the content over time. The course schedule is displayed in Appendix D.

Experimental Design

An alternating treatments design (Kazdin, 1982) was used to compare the three treatment conditions (pairs, small groups, large groups). According to What Works Clearinghouse (2008), requirements for strong or moderate causal evidence when comparing two conditions include at least 3 demonstrations of an effect in the same direction, no clear effects in the opposite direction and the overall mean levels for the two interventions clearly demonstrating a visual effect. The independent variable was discussion group size (pairs, small groups, or large groups) while the dependent variables were student performance (mean quiz scores and mean scores on a cumulative final) and social validity data (student satisfaction following each interteach group and student preference at the end of the quarter) derived from the teaching record and an end of the quarter questionnaire.

Independent Variable. Group size was the primary independent variable in this study. Students alternated between pairs, small groups, and large groups depending on the random assignment of the class. Each student was randomly assigned to the discussion groups. Other than the size of discussion group, the implementation of interteaching did not vary across conditions.

Dependent Variables. Student performance was measured as mean quiz scores (%), and mean scores on a cumulative final (%). Quiz scores were analyzed on an alternating treatment design graph across each of the treatment groups.

Social validity (satisfaction and preference) was measured when analyzing student responses to the teaching record and an end of the quarter questionnaire. Students were asked to provide information pertaining to the quality of their interteaching group (Likert Scale 1-5) each week as well as their overall preference when the quarter ended.

Interobserver Agreement

For each exam, a computerized scoring/assessment device was used to assess students each week. To determine interobserver agreement (IOA), the professor and teaching assistant independently completed 20% of the exam questions to establish average IOA across the exams. The number of questions that were identically graded by the instructor and teaching assistant was divided by the total number of items analyzed multiplied by 100. When the TA and professor disagreed, they discussed the item and came to agreement on the final score.

Procedural Fidelity

The goal of assessing procedural fidelity was to ensure that each treatment condition included each component of the interteach process. To achieve this goal, procedural fidelity was measured through the use of a procedural fidelity checklist. Each component of the interteach process (prep guide, discussion groups, teaching record, clarifying lecture, and quiz) was identified and checked off during each class by the TA on 33% of classes (each condition was assessed for procedural integrity at least twice) to ensure that the interteaching procedure was completed with fidelity. During each of these checks, all components occurred with 100% accuracy. In the event that a specific component did not occur, the TA would either remind the

instructor in the moment or the data would not be reflected in the final analysis. However, this did not occur.

Results

Quiz Scores

Students' scores on each of the quiz scores was a primary dependent variable. Figure 1 shows that quiz scores following group sessions consisting of pairs or small groups show relatively no difference. Quiz scores following large groups appear to be slightly below that of pairs and small groups. The cumulative final served as a way to assess generalization of knowledge over time. Each question on the cumulative final was taken specifically out of prep guides that were used during group discussion (17 questions from prep guides used during pair groups, 17 questions from prep guides used during small groups, and 17 questions from prep guides during large groups). The remaining 15 questions were related to questions covering the final assigned readings and ethics in behavior analysis. Figure 2 shows that students scored higher on quizzes when they were grouped in pairs and small groups (M= 9.24 and 9.23, respectively) compared to large groups *(*M = 8.90).

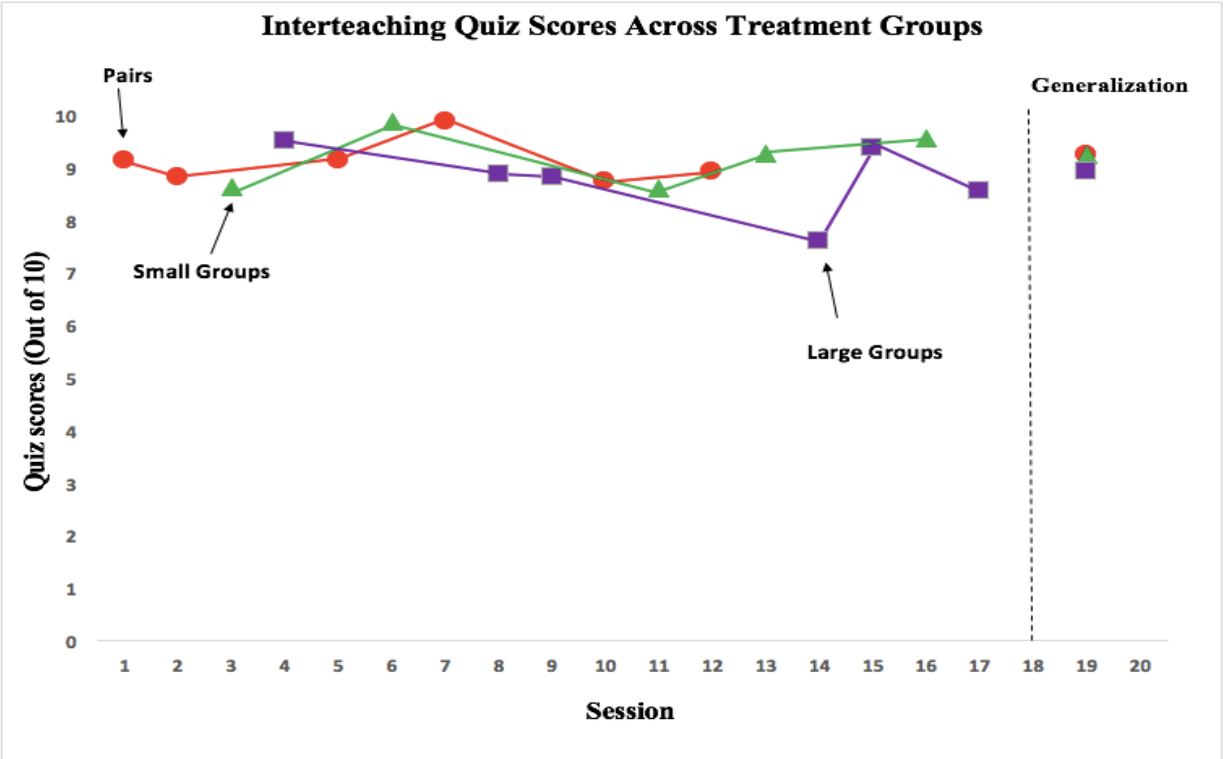

Figure 1. Interteaching Quiz Scores Across Treatment Groups.

Mean scores across all treatment conditions showed very little difference, however students scored highest on quizzes following small groups ($M = 9.19$) compared to pairs and large groups ($M = 9.12$ and 8.80, respectively). While this difference is small, it is a relative important distinction given the recommendations provided by previous research on the subject.

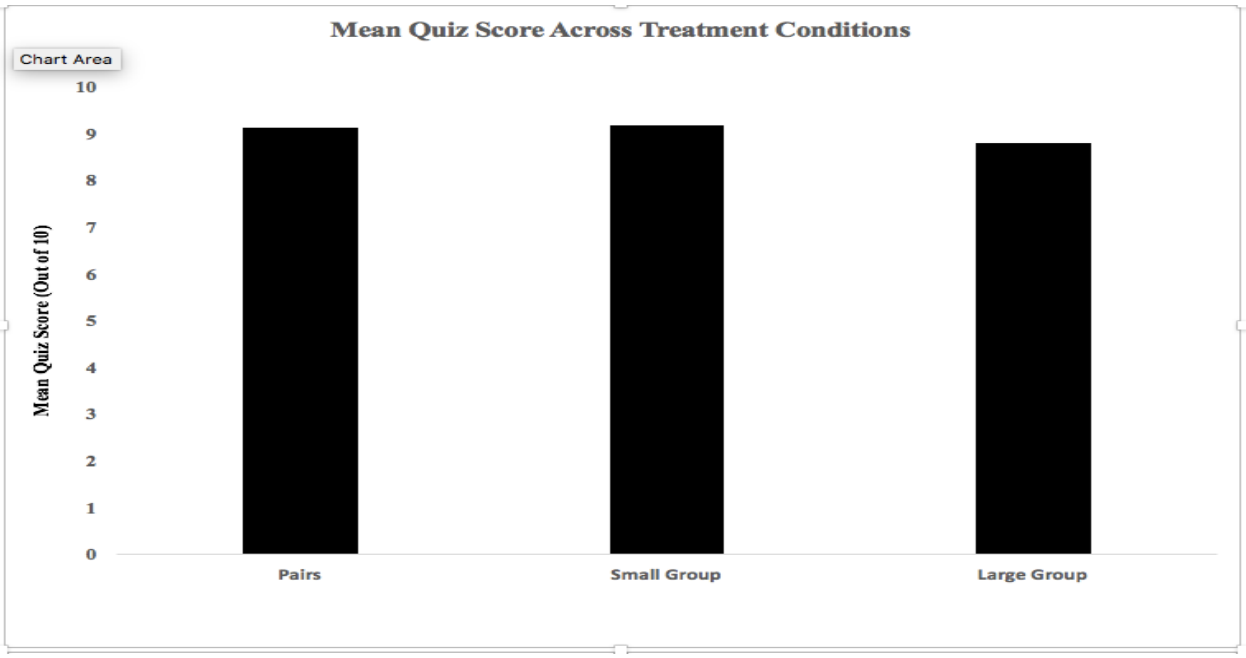

Figure 2. Mean Quiz Scores Across Treatment Conditions

Satisfaction and Preference

Following each group, students were asked to complete a teaching record which provided important information related to comprehension of the assigned material as well as student satisfaction related to the quality of their group. Students were asked to rate their level of satisfaction on a scale of 1 (highly not satisfied) to 5 (highly satisfied). Figure 3 shows that student's showed much higher rates of satisfaction with the quality of their group session following small group sessions (M = 4.81) compared to pairs and large group sessions (M = 3.38 and 3.06, respectively).

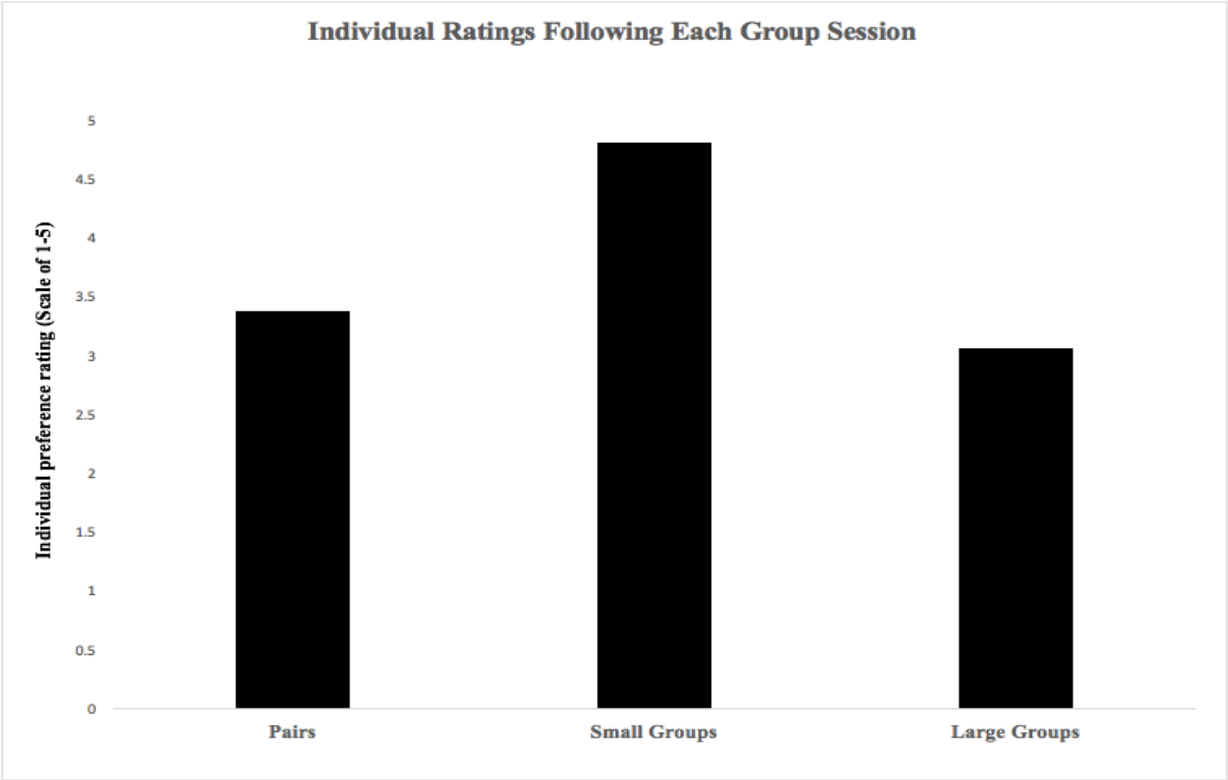

Figure 3. Individual Ratings Following Each Group Session

At the end of the course, students filled out a survey asking them questions related to group size preference as well as the pros and cons related to each. When specifically choosing the size group they prefer to meet in, students overwhelming chose small groups (17 students total) compared to pairs (1 student) and large groups (1 student). See Figure 4.

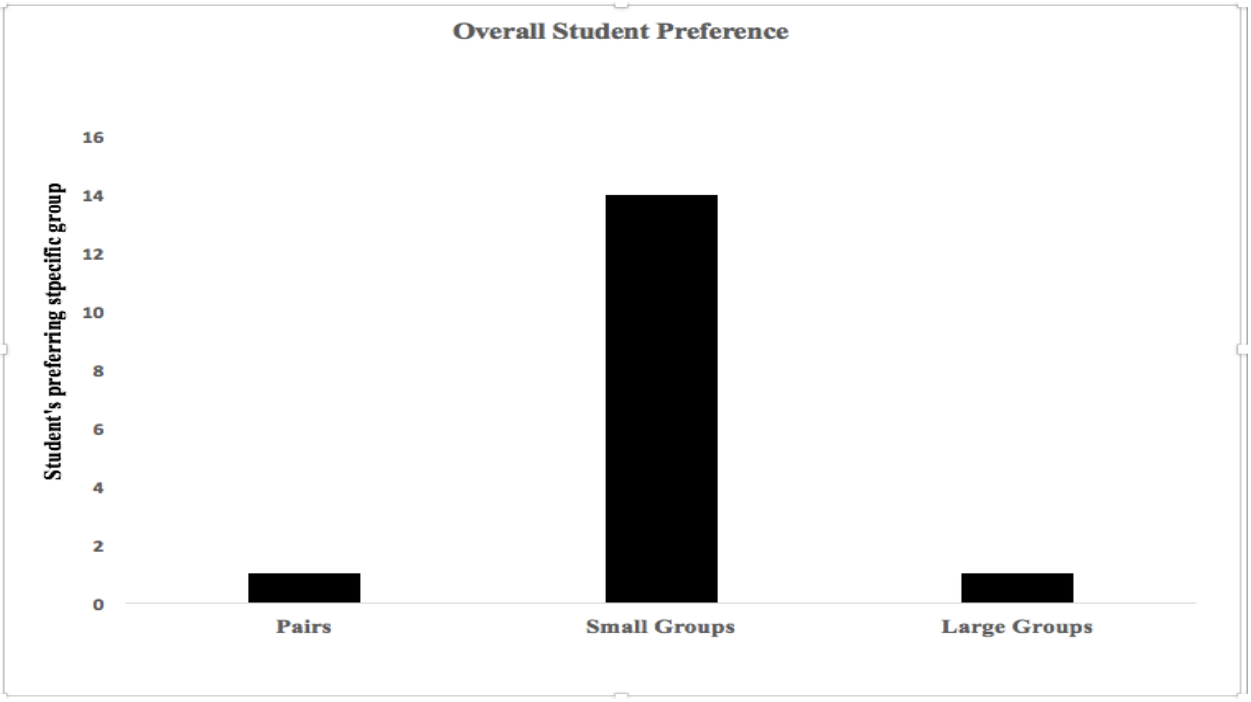

Figure 4. Overall Student Preference for Group Size

Social Validity

When asked to provide information related to the pros and cons related to meeting in pairs, students provided a lot of useful information. Six students specifically mentioned "intimacy" as a primary positive aspect of meeting in pairs. Five students identified increased time to discuss the material due to a lower number of students participating in the discussion as a pro related to working in pairs. Other comments related to positive aspects of pair groups include comfortability (2 students), less structured (2 students), increased ability to go "in-depth" with the material (2 students), and more support compared to other group sizes. Together, these comments specifically identified the affordances that meeting in pairs may provide related to increased intimacy and feelings of comfortability to go deeper into the material compared to other groups.

When asked to provide information related to cons of meeting in pairs, 10 students wrote that they felt too "reliant" on their partner when it came to helping them answer questions, clarifying information, or helping them to better understand the material. Five students wrote that a major con to meeting in pairs is that there are less varied opinions/perspectives. Three students wrote about issues related to unwilling partner participants. Specifically, if a partner was non-participative or did not read the assigned readings, it impacted the quality of the group. Lastly, students wrote that there were not enough people to contribute, it took too long to get through the material, and that it was easier to deviate from the topic.

Students wrote a variety of perspectives related to meeting in small groups as well. Of all the groups, small groups had the largest number of pros identified (9 in total). Within small groups students wrote that the group was better prepared to identify more examples related to course material (8 students), still intimate enough for students to feel comfortable to contribute (7 students), and had a variety of opinions and perspectives which led to better understanding of the source material (6 students). Five students mentioned that it was a good number related to facilitating participation and that it was the "best of both worlds" (large enough to get a decent variety of opinions/perspectives, but not too large where students shy away from participation). Students mentioned that the discussion was of high quality (4 students), that there was more reliability related to the quality of answers provided by the group (3 students), and that the group was efficient with how it progressed through the preparation guides. One student felt that groups of 4 was better at facilitating in-depth discussions related to the material.

When asked to identify specific cons to small groups, students gave the least amount of answers (3 total). One student mentioned that there were less examples provided related to source material. Another student mentioned the concept of "group stars" where a few select

students may hog the spotlight during group meetings making it less inclusive for timid participants. Two students wrote that even with four students, they may still need help with source material.

Students' responses to identify the pros and cons of large groups were quite varied. Eight different pros were identified. By far, the most identified advantage to groups of 8 students were the varied answers, opinions, and perspectives when meeting within such a large group of students (10 students). Following the identification of this advantage, the number of responses dropped off. Four students identified an increase in examples related to the material. Three students mention that there was increased reliability in the answers and that they were able to get through the entire preparation guide with confidence that they have the majority of correct answers. Two students mentioned that it was less awkward than small groups and pairs since they didn't have to participate in the same way that the other group's natural contingencies may have forced them to.

As far as disadvantages identified by students toward a specific group, large groups had the greatest number (ten in total). Ten students mentioned that there was so much variation that it led to confusion. Another ten students wrote that it was too large for any meaningful discussion. Five students mentioned that it was less intimate compared to smaller groups. Four students specifically mentioned the "group stars" phenomenon where certain individuals spoke more than the group members felt required. Other disadvantages identified included: hard to focus (4 students), the group moving too quickly through the material (3 students), less participation for select students (3 students), too much reliance on others to contribute (2 students), uncomfortable (2 students), and too disorganized (1 student).

Discussion

The current study investigated the effects of interteaching group size on student performance (quiz scores and a final exam) and student preference during an undergraduate course teaching the principles of applied behavior analysis. Results showed that group size had very little impact on student performance across quizzes and a final exam. It is important to note, however, that quiz scores were high across all three conditions (M = 9.03), and that there was a slight improvement in test scores during the small group condition (M = 9.19). When examining preference, on the other hand, group size appears to have a strong effect. Students overwhelmingly preferred small groups over pairs and large groups. Within the broad context of the interteaching literature, these findings are relevant when it comes to course planning as well as meeting the demands of larger course enrollments and alternate delivery formats (e.g., online teaching).

Boyce and Hineline's (2002) article states that groups of two are the most optimal because total student engagement and involvement is maintained by natural contingencies of social interaction. Truelove's (2013) study on group size found that no significant differences exist between groups of two and four across 6 unit exams, a cumulative final, and in the total number of points earned across the semester. The authors mention that given the fact that performance may deteriorate with increased group size, additional research is needed to determine if this is true. The present study found similar results in that pairs and small groups had no differences in student performance, while students did slightly less well on assessments in large groups. Students in this study also stated a strong preference for small groups over pairs and large groups. These findings have various implications.

Implications

According to research conducted by Mitchell et al. (2016), following the recession of 2007-2009 record breaking declines in state revenues led to increased state funding cuts to higher education funding. As a result, public institutions of higher learning have been forced to reduce the quality and availability of many of their academic offerings. One evidence of this is increased class sizes. Specifically, between the years of 2008 and 2014, the number of full time equivalent instructional staff at public colleges and universities grew by about 7% while the number of students at these institutions grew by 8.6%. In other words, the number of students per faculty member is rising nationwide. Given this information, instructors should find evidence-based teaching that can be effective even with increases in class size.

If an instructor is expected to teach a classroom with 40 students using the current guidelines for interteaching, they would be expected to intermittently check in with 20 different pairs of students. The results of the current study suggest that splitting into groups of 4-5 will have little to no impact on student performance and my even boost student satisfaction compared to be grouped in pairs. Further, this would result in the instructor being able to float between 10 groups rather than 20, resulting in increased instructor access as well as reinforcement for correct answers and staying on topic. Further research should be conducted to confirm these results so that instructors have greater flexibility to alter group size to suit the needs of their specific classroom.

These findings can be especially relevant in field of online learning. An increasing number of students are joining the online community (National Center for Statistics, 2014) and currently there are no studies examining the effects of interteaching in the field of online learning. When conducting interteaching online, instructors must group students into pairs and

bounce from chat room to chat room to check in on each group. If the class size is any number higher than 20, investigators must bounce between at least 10 online chat rooms to meet the requirements of interteach. Due to this requirement of pair groups within the interteach process, instructors may be reluctant to use interteaching online. However, the present findings may alleviate some of these difficulties by allowing larger group size.

Another important implication that the current study finds is related to student satisfaction and preference. Student satisfaction is a well-researched topic in the field of academic settings (Dhaqane & Afrah, 2016). Research has shown that higher levels of satisfaction leads to higher levels of relative performance across a variety of domains including participation, academic performance, student retention, and longevity in the field, (Biner, Barone, Welsh, & Dean, 1997; Dong & Lucey, 2013; Douglas, McClelland, & Davies, 2008; Tucker, Pegden, & Yorke, 2012; Wiers-Jenssen, Stensaker, & Grogaard, 2002). With this research in mind, it is of utmost important for instructors to regularly assess how specific teaching methodologies may affect student satisfaction and preference.

Limitations

This study had a number of limitations that are pertinent to consider. First, the relative size of our student population was small (19 students). It is recommended that further studies attempt to replicate and expand on these findings in order to provide increased evidence that the claims provided in this paper have weight.

Another limitation of this study is related to the randomization process. When randomizing the classes in order to provide each with a pre-identified group number, the majority of classes where pairs groups were administered took place toward the beginning of the course rather than spread out evenly. Given that the course may progressively become more

difficult as the subject matter increases in scope and complexity, classes that utilized pair groups may have been exposed to subject matter that is easier to understand. If researchers attempt to replicate this study, it is recommended that the groups are spread out evenly across the class so that they are exposed to a variety of material with varying complexity.

Lastly, students did not know what the group size would be prior to class. It is possible that students knowing the group size ahead of time may affect the work they put into the preparation guide or if they complete it at all. Given that students did not know the assignment for each class, this may have affected the results.

Future research

There are a number of areas where future research could prove beneficial considering these findings. First, replication across larger student populations could help increase the reliability of these findings. Second, replication within the field of online learning could provide instructors with the confidence that they could use interteaching in their online classrooms. Also, it is also suggested that future research investigate instructor perspectives on the interteach process to assist in the uptake and continued growth of interteaching within the educational community. Lastly, it would beneficial to study the effects of group size when the students know ahead of time what the group assignment will be. This information could shed light on the effects of group size could be on performance and quality of groups when the course utilizes the same group size throughout an entire quarter or semester.

References

Arntzen, E., & Hoium, K. (2010). On the effectiveness of interteaching. *The Behavior Analyst Today*, *11*(3), 155.

Biner, P., Barone, N., Welsh, K., & Dean, R. (1997). Relative academic performance and its relation to facet and overall satisfaction with interactive telecourses. *Distance education*, *18*(2), 318-326.

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*(2), 215.

Brown, T. W., Killingsworth, K., & Alavosius, M. P. (2014). Interteaching: An Evidence-Based Approach to Instruction. *International Journal of Teaching and Learning in Higher Education*, *26*(1), 132-139.

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.

Cullen, J. G. (2011) The writing skills course as an introduction to critical practice for larger business under-graduate classes. *International Journal of Management Education* 9(4): 25–38.

Delquadri, J., Greenwood, C. R., Whorton, D., Carta, J. J., & Hall, R. V. (1986). Class wide peer tutoring. *Exceptional Children*, *52*(6), 535-542.

Dong, Y., & Lucey, A. (2013). Relationships between student satisfaction and assessment grades in a first-year engineering unit. *Teaching and Learning Forum, 2013*

Douglas, J., McClelland, R., & Davies, J. (2008). The development of a conceptual model of

student satisfaction with their experience in higher education. *Quality Assurance in Education*, *16*(1), 19-35.

Dhaqane, M. K., & Afrah, N. A. (2016). Satisfaction of Students and Academic Performance in Benadir University. *Journal of Education and Practice*, *7*(24), 59-63.

Goodlad, S., & Hirst, B. (1989). *Peer Tutoring. A Guide to Learning by Teaching*. Nichols Publishing, PO Box 96, New York, NY 10024.

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*(1), 31-35.

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

Kulik, J. A., Carmichael, K., & Kulik, C. L. (1974). The Keller Plan in Science Teaching: An individually paced, student-tutored, and mastery-oriented instructional method is evaluated. *Science (New York, NY)*, *183*(4123), 379-383.

Lindsley, O. R. (1991). Precision teaching's unique legacy from BF Skinner. *Journal of Behavioral Education*, *1*(2), 253-266.

Malott, R., & Shane, J. T. (2015). *Principles of behavior*. Psychology Press.

Mitchell, M., Leachman, M., & Masterson, K. (2016). Funding Down, Tuition Up. *Center on Budget and Policy Priorities.* Retrieved from http://www.cbpp.org/research/state-budget-and-tax/funding-down-tuition-up

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.

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*(4), 937-941.

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*(3-4), 27.

Saville, B. K., & Zinn, T. E. (2009). Interteaching: The effects of quality points on exam scores. *Journal of Applied Behavior Analysis*, *42*(2), 369-374.

Saville, B. K., Zinn, T. E., & Elliott, M. P. (2005). Interteaching versus traditional methods of instruction: A preliminary analysis. *Teaching of Psychology*, *32*(3), 161-163.

Saville, B. K., Zinn, T. E., Neef, N. A., Norman, R. V., & Ferreri, S. J. (2006). A comparison of interteaching and lecture in the college classroom. *Journal of applied behavior analysis*, *39*(1), 49-61.

Schwartz, I., Ashmun, J., McBride, B.J., Scott, & Sandall, S. (2017). *The Data Model for Teaching Preschoolers with Autism*. Brookes Publishing Co.

Scoboria, A., Sirois, F. M., & Pascual-Leone, A. (2009). Using Interteaching to Enhance Student Engagement and Learning. *Collected Essays on Learning and Teaching*, *2*, 83-88.

Slavin, R. E. (1980). Cooperative learning. *Review of educational research*, *50*(2), 315-342.

Soldner, J. L. (2014). Enhancing the impact of quality points in interteaching. *Journal of the Scholarship of Teaching and Learning*, *14*(5), 1-11.

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*(2), 23-30.

Tsui, M. (2010). Interteaching Students as Teachers in Lower-Division Sociology

Courses. *Teaching Sociology*, *38*(1), 28-34.

Tucker, B., Pegden J.-A. & Yorke, J. (2012). Outcomes and evaluations: Is there a relationship between indicators of student success and student evaluations of learning? In Brown, N., Jones, S. M. & Adam, A. (Eds.) *Research and Development in Higher Education: Connections in Higher Education*. Proceedings of the 35th HERDSA Annual International Conference, 2-5 July 2012.

Clearinghouse, W. W. (2008). Procedures and standards handbook (Version 3.0). Retrieved June, 6, 2011.

Wiers-Jenssen, J., Stensaker, B. R., & Grogaard, J. B. (2002). Student satisfaction: Towards an empirical deconstruction of the concept. *Quality in higher education*, *8*(2), 183-195.

Wood, D. F. (2003). Problem based learning. *BMJ: British Medical Journal*, *326*(7384), 328.

Appendix A

Behavior Analysis Certification Board's (BACB) required Registered Behavior Technician (RBT) Task List

A. Measurement
A-01 Prepare for data collection **A-02** Implement continuous measurement procedures (e.g. frequency, duration). **A-03** Implement discontinuous measurement procedures (e.g. partial & whole interval, momentary time sampling) **A-04** Implement permanent product recording procedures **A-05** Enter data and update graphs
B. Assessment
B-01 Describe the behavior and environment in observable and measurable terms **B-02** Conduct preference Assessments **B-03** Assist with individualized assessment procedures (e.g. curriculum-based, developmental social skills). **B-04** Assist with functional assessment procedures
C. Skill Acquisition
C-01 Identify the essential components of a written skill acquisition plan. **C-02** Prepare for the session as required by the skill acquisition plan. **C-03** Use contingencies of reinforcement (e.g., conditioned/unconditioned reinforcement, continuous/intermittent schedules). **C-04** Implement discrete-trial teaching procedures. **C-05** Implement naturalistic teaching procedures (e.g., incidental teaching). **C-06** Implement task analyzed chaining procedures. **C-07** Implement discrimination training. **C-08** Implement stimulus control transfer procedures. **C-09** Implement stimulus fading procedures. **C-10** Implement prompt and prompt fading procedures. **C-11** Implement generalization and maintenance procedures.

C-12 Assist with the training of stakeholders (e.g., family, caregivers, other professionals)
D. Behavior Reduction
D-01 Identify the essential components of a written behavior reduction plan. **D-02** Describe common functions of behavior. **D-03** Implement interventions based on modification of antecedents such as motivating/ establishing operations and discriminative stimuli. **D-04** Implement differential reinforcement procedures (e.g., DRA, DRO). **D-05** Implement extinction procedures. **D-06** Implement crisis/emergency procedures according to protocol
E. Documentation and Reporting
E-01 Report other variables that might affect the client (e.g., illness, relocation, medication). **E-02** Generate objective session notes by describing what occurred during sessions. **E-03** Effectively communicate with supervisor. **E-04** Comply with applicable legal, regulatory and workplace reporting requirements (e.g., mandatory abuse and neglect reporting). **E-05** Comply with applicable legal, regulatory and workplace requirements for data collection, storage and transportation.
F. Professional Conduct and Scope of Practice
F-01 Describe the role of the RBT in the service delivery system. **F-02** Respond appropriately to feedback and maintain or improve performance accordingly. **F-03** Communicate with stakeholders (e.g., family, caregivers, other professionals) as authorized. **F-04** Maintain professional boundaries (e.g., avoid dual relationships, conflicts of interest, social media contacts). **F-05** Maintain client dignity.

Appendix B

Preparation Guide Example

Reading: Chapter 22 and 23

Directions: Please read the assigned readings first, then answer the questions below. You will meet in groups during your next class to discuss your answers to the questions below.

Chapter 22

1. What does anorexia nervosa mean?
2. Describe and analyze the intervention for Mary's anorexia.
3. What part of the intervention seems to be based on simple reinforcement and what part seems to involve reinforcers that are too delayed to be useful in simple reinforcement
4. What were the results of the intervention?
5. For each of the following concepts: define it and give an example:
    a. Rule
    b. Rule control
    c. Rule-governed behavior
    d. Contingency control
    e. Contingency-governed behavior (contingency-shaped behavior).
6. Give an example where behavior is controlled by the promise of the reinforcer rather than the delivery.
7. Give an example where an apparent delayed analog to reinforcement is really a delayed analog to discriminated avoidance.
    a. Include a discussion of the role of the deadline
    b. Be able to diagram your example
8. Give an example where an apparent delayed analog to reinforcement is really a delayed analog to nondiscriminated avoidance, with the time before the deadline merely being the time when there's an opportunity to respond.
    a. Include a discussion of the role of the deadline
    b. Be able to diagram your example
9. Define and give an example of each of the following concepts:
    a. Direct-acting contingency
    b. Indirect-acting contingency
    c. Ineffective contingency
10. Construct the contingency tree
11. Rule governed analog to reinforcement by the presentation of a reinforcer- state it and give an example, and explain how your example fits the definition
12. Give an example showing how such an analog can indirectly control behavior
13. Describe the traditional and behavioral parenting techniques involved with Wayne.
14. Diagram and label the contingency Wayne's mother used during the behavioral intervention.
15. Give an example of a statement of a rule, a related instructional statement, and a related statement of a request.
16. Present an argument for the necessity of analyzing some sorts of behavior in terms of rule-governed behavior.

Chapter 23

1. Define and give an example of the following:
    a. Task analysis
    b. The process vs. product general rule
    c. Feedback
2. Diagram and explain the role of delayed feedback in a behavior-analytic coaching procedure.
3. For each of the following areas, give an example of research showing the successful use of a rule-governed analog to avoidance of the loss of the opportunity to receive a reinforcer. Explain how your examples are best understood in terms of this contingency and diagram the contingencies
    a. Education
    b. Medicine
    c. Carpooling
    d. Litter control
    e. Dental care
4. What two common errors do behavior analysts make concerning rule-governed analogs?
5. Give an example of each type and explain why they are errors.
6. Give three of our examples of how information was not enough to get people to act in their best interests and the best interests of the community.
7. When should you put contingencies on behavior leading to the accomplishment and not just on the accomplishment?
8. List the four issues to address when putting contingencies on the process in addition to the product.
9. Multiple baseline design-define it
    a. Why is it useful?
    b. Give an example of a multiple baseline across groups.
    c. And one across behaviors.
10. Covert behavior- define it and give an example
11. Reinforcement of covert behavior- give an example
12. The principle of shifting from rule control to contingency control- state it, give three examples, and explain how your examples fit the principle.

Appendix C

Teaching Record

Date: _____

Duration of Discussion: _____

How many students were in your group including yourself? _____

Was there sufficient time provided: Yes_____ NO_____

	Very Easy	Easy	Neutral	Difficult	Very Difficult
How difficult was the material covered in group?	1	2	3	4	5

What topic(s) did you find difficult to understand? What about the topic did you not understand?

_____

_____

What topic(s) would you like clarification on during the next class lecture?

_____

_____

	Highly not satisfied	Not Satisfied	Neutral	Satisfied	Highly Satisfied
Rate your level of satisfaction with your discussion group.	1	2	3	4	5

What did you like or not like about your group during this class period?

_____

_____

Appendix D

Course Schedule

Date	Study Guide	Lecture	Quiz	Group Discussion	Teaching Record	Group #
1/3	x	Intro	x	x	general	N/a
1/5	Chap. 1	Practice	Practice	Chap. 1	Chap. 1	2
1/10	Chap. 2&3	Chap. 1	Chap. 1	Chap. 2&3	Chap. 2&3	2
1/12	Chap. 4&5	Chap. 2&3	Chap. 2&3	Chap. 4&5	Chap. 4&5	4
1/17	Chap. 6	Chap. 4&5	Chap. 4&5	Chap. 6	Chap. 6	8
1/19	Chap. 7	Chap. 6	Chap. 6	Chap. 7	Chap. 7	2
1/24	Chap. 8 & 20	Chap. 7	Chap. 7	Chap. 8 & 20	Chap. 8 & 20	4
1/26	Chap. 9 & 10	Chap. 8 & 20	Chap. 8 & 20	Chap. 9 & 10	Chap. 9 & 10	2
1/31	Chap. 11	Chap. 9 & 10	Chap. 9 & 10	Chap. 11	Chap. 11	8
2/2	Chap. 12	Chap. 11	Chap. 11	Chap. 12	Chap. 12	8
2/7	Chap. 13	Chap. 12	Chap. 12	Chap. 13	Chap. 13	2
2/9	Chap. 14 and Project DATA Chap.	Chap. 13	Chap. 13	Chap. 14 and Project DATA Chap.	Chap. 14 and Project DATA Chap.	4
2/14	Chap. 15 & 16	Chap. 14 and Project DATA chapter	Chap. 14 and Project DATA Chap.	Chap. 15& 16	Chap 15& 16	2
2/16	Project DATA Chap.	Chap. 15& 16	Chap. 15 & 16	Project DATA Chap.	Project DATA Chap.	4
2/21	Chap. 17 & 18	Project DATA Chap.	Project DATA Chap.	Chap. 17 & 18	Chap. 17 & 18	8
2/23	Chap. 19 & 21	Chap. 17 & 18	Chap. 17 & 18	Chap. 19 & 21	Chap. 19 & 21	8
2/28	Chap. 22 & 23	Chap. 19 & 21	Chap. 19 & 21	Chap. 22 & 23	Chap. 22 & 23	4
3/2	Chap. 27 & 28	Chap. 22 & 23	Chap. 22 & 23	Chap. 27 & 28	Chap. 27 & 28	8
3/7	Ethics Readings	Chap. 27 & 28	Chap. 27 & 28	Ethics Readings	Ethics Readings	4

3/9	x	Ethics Readings	Ethics Readings	x	x	N/A