8-2012

Using Cooperative Learning Groups to Enhance Classroom Participation in a Large Undergraduate Course

Cora Marie Taylor
ctaylo63@utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Cora Marie Taylor entitled "Using Cooperative Learning Groups to Enhance Classroom Participation in a Large Undergraduate Course." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in School Psychology.

Robert L. Williams, Major Professor

We have read this dissertation and recommend its acceptance:

Sherry K. Bain, Sherry M. Bell, David F. Cihak

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Using Cooperative Learning Groups to Enhance Classroom Participation in a Large Undergraduate Course

A Dissertation Presented for the Doctor of Philosophy Degree
The University of Tennessee, Knoxville

Cora Marie Taylor

August 2012

To Megan – my first teacher.

Acknowledgements

First and foremost, I would like to extend my thanks to Dr. Robert Williams for his teaching and mentoring over the course of my graduate education and throughout this dissertation process. He has been an invaluable source of guidance and my work with him during my time at the University of Tennessee has truly helped to shape my skills as a researcher and teacher. I am indebted to the other members of my committee, Dr. Sherry Bain, Dr. Sherry Mee Bell, and Dr. David Cihak whose feedback has been essential to the success of this dissertation. My gratitude is also extended to Carolyn Blondin, Bethany Forbes, Charles Galyon, and Jared Yaw who spent valuable time helping with data collection. I also owe great thanks to Mike O'Neil, Eric Heidel, and Dr. Dennis Ciancio for their statistical guidance.

An important thank you is owed to all of the faculty with whom I have had the great pleasure of working, especially Dr. P. Niels Christensen who first encouraged me to pursue graduate education and provided me with the tools to be successful, as well as Drs. Chris Skinner and Brian Wilhoit who have given me priceless instruction and advice to help shape my practice in the field of school psychology. In addition, I am thankful to my mentor, Dr. Daniel McCleary, for serving as a guide not only through the dissertation process, but throughout my entire graduate program. Finally, I would like to extend a special thank you to my family and friends who have offered endless support from near and far.

Abstract

This research study focused on the use of cooperative-learning groups to facilitate classroom participation in a large undergraduate course. Data were collected in three sections of an Educational Psychology course ($n \approx 56$ per section). At the conclusion of the first class unit (in which no credit for participation was available), students were assigned to cooperative groups based on their participation. Each group consisted of five to six students whose participation in the first unit ranged from low to high. At the conclusion of each remaining unit (total of four units), two days were randomly selected for individual participation credit. Students could receive up to five points for each selected day (three points for the first comment and two additional points for a second comment). In addition to the possibility of individual credit in these four units, students could receive bonus credit in two of the four credit units if every present member of their group participated at least once in class on each day selected for individual credit. Group members received 5 bonus points for each day their group met this criterion, allowing for 10 bonus points available for each unit selected for group-plus-individual credit.

Data from the study were entered into an SPSS database, and results were analyzed through visual inspection of graphs and three-way mixed designs. Analyses showed that the first application of credit for participation was more effective than the second application of credit for participation. Additionally, the individual-plus-group credit contingency produced greater participation than individual credit alone. In addition to assessing the effect of the credit contingencies on class-wide participation, I examined the differential effects of the credit contingencies on initially-low participants. In two sections of the course, individual-plus-group credit was consistently more effective in increasing participation of initially-low participants

than was individual credit alone. However, the initially-low participants in one section of the course were unresponsive to all applications of both types of credit. Critical thinking scores were analyzed as a potential covariate for participation. The group-bonus contingencies appeared to have minimal effects on exam performance across cooperative teams in the three sections.

Table of Contents

List of Tables

List of Figures

Chapter I

Introduction and Literature Review

Classroom participation in college courses has been an important and widely-researched topic in educational literature. Research on typical classroom participation has indicated that only a few students (approximately 20% of the students in attendance) account for the majority of class engagement (Howard, Short, & Clark, 1996; Fritschner, 2000). More specifically, in large introductory college courses, only about six students participate during each class period, and only three of these students participate more than once during a class session (Fritschner). Students tend to rank their peers as being very uninvolved in class discussion, despite ranking themselves as highly involved (Burchfield & Sappington, 1999). Further research has indicated that students tend to see their peers as less involved in classroom discussion than does the professor (Fassinger, 1995a).

Fassinger (1995a) determined a variety of student factors that seem to greatly affect amount of participation in college classes. The most critical student factor appears to be confidence. Not surprisingly, students who report higher levels of self-confidence are more inclined to participate during discussion. Furthermore, it appears that the emotional climate of the classroom (as set by other students, not just the professor) heavily influences amount of participation in class. This finding is counter to a more traditional notion that teachers primarily create the climate of participation or non-participation. Rather, it seems that the design of the class (primarily discussion-based or lecture-based) is more important in fostering participation than the interpersonal style of the professor (Fassinger, 1995b). Additional analysis has demonstrated that classrooms with positive group dynamics (e.g., interdependence among

students, more cooperation than competition, and support of others' comments) are more likely to encourage students to participate. While these positive classroom dynamics can increase participation, they can also have negative effects on participation, including avoidance of more controversial topics, which may be integral to course content and discussion (Fassinger, 2000).

Discussion on grading methods of class participation has been an important part of research on class participation. Although Bean and Peterson (1998) acknowledged the disadvantages to grading class participation (e.g., grading tends to be subjective, shy students can be at a disadvantage, and students are generally not given instruction on how to improve their participation), they believe that finding appropriate procedures for grading participation can greatly improve classes by sending signals to students that participation is highly important and can lead to better student preparation for class sessions. The authors suggest a grading rubric for participation that would create a more objective measure for teachers to judge the quality of students' class preparation and participation.

Despite the admitted difficulty of monitoring and grading classroom participation in college courses, Boniecki and Moore (2003) attempted to create a manageable token economy system to reinforce participation in an undergraduate introductory psychology classroom. In their system, students received a small token each time they voluntarily and correctly answered a question posed by the instructor. At the end of each class period, the students could exchange their tokens for one extra credit point on their next exam. Any tokens that were not exchanged at the end of the class period were void and could not be used at any future date. Overall, each token was worth approximately 0.25% of the final grade. As expected by the researchers, students raised their hands more when tokens were available than during the baseline and

removal of tokens. Furthermore, students raised their hands faster after a question was posed by the instructor, and students engaged in more spontaneous participation (e.g. asked more questions) during the token phase than either the baseline or removal phases.

In contrast to the above study, Hodge and Nelson (1991) used a method of differential reinforcement to improve classroom participation. The researchers, with a small class of 16 undergraduate psychology students, identified students who needed to increase participation and several dominant students who needed to decrease their participation. Using the symbol of a plus mark on the board next to student names, the instructor visually reinforced students for engaging in participatory behavior for low or non-participants, and quiet or reserved behavior for dominant students. Based on observations by the instructor and fellow students (at the conclusion of the study), most members of the class noticed an improved balance of participation across class members. According to the instructor, only one highly reticent member of the class appeared to be unaffected by the procedure.

Connor-Greene (2005) has encouraged professors to give students daily homework that allows students to carefully examine the readings in order to prepare for classroom discussion. She recommends an approach called Question, Quotation, and Talking Points (QQTP) to reflect on the day's readings. The author reported that the QQTP system allowed her to provide objective grades (based on completion of this daily activity related to the readings), while simultaneously encouraging students to engage in the preparation required to be beneficial, active participants in class.

Cooperative-Learning Groups

One method that educators have implemented to improve academic performance within their classrooms has been the use of cooperative-learning groups. While the relationship between

the use of cooperative-learning group contingencies and classroom participation has not been directly studied, cooperative-learning groups have been shown to improve academic outcomes in most cases (Slavin, 1991). Furthermore, cooperative-learning groups have produced positive outcomes in many areas other than academic success, including improvements in behavior, social acceptance, and attendance (Slavin, 2011).

Research on cooperative learning has included the investigation of different types of cooperative-learning groups to determine which structure is the most effective. Slavin (2011) outlined several models of cooperative groups that have demonstrated positive academic achievement in schools. The first type of cooperative group is called Student Teams-Achievement Divisions (STAD). STAD groups generally have four to five members, and the group is organized to mirror the makeup of the class as a whole in terms of achievement, race, and gender, as much as possible. Thus, each group would likely contain at least one high achiever, at least one low achiever, and two to three average achievers in the class. After class material has been presented, students spend time reviewing material and studying with their teams. Next, students individually take a short quiz on the assigned material, and receive an individual grade for their performance. In addition to their individual grade, teams earn points for each group member's improved performance and receive team recognition for the composite improvement of the group.

Teams-Games-Tournaments (TGT) is another method of organizing cooperative groups. As in STAD, teams are generally made up of four to five members, and each team is representative of the class as a whole. After new material is presented, class members have a chance to study with their teams. After a study period, students break into tournament teams, which, unlike the study teams, are not representative of the entire class. Instead, these groups are

made up of students who have similar levels of achievement. Thus, a low achiever would go to a tournament team with other low-achieving students and a high achiever would go to a tournament team with other high-achieving students. While in tournament teams, students quiz each other on the new material and compete for points to bring back to their study teams. At the conclusion of the tournaments, each team adds together all points that the members won during tournaments, and teams can earn rewards and recognition for teams who have improved their performance or maintained high levels of performance from previous tournaments (Slavin, 1991; Slavin, 2011).

Jigsaw II is another prominent model of cooperative learning in the classroom. In this model of collaboration, students are organized both as study teams and expert groups. Each individual in study teams is responsible for learning information on an assigned topic. After a general review of all informational assignments with study-team members, each team sends each member to an expert group that will address the team member's assigned topic. Members of the expert groups help one another become experts on the assigned topic by using guided questions and worksheets. Once students have spent time learning with expert groups, they go back into teams, with each team having one member reporting back from each expert group. Team members are responsible for teaching one another the information they learned while in the expert group. After teams have exchanged information and studied all of the material, students take a quiz, and members can earn points for their teams based on individual academic improvement (Slavin, 1991).

Although all of the research on the structure of cooperative-learning groups helps educators learn new ways to promote student achievement and cooperation, group structure alone does not fully explain the effectiveness of cooperative-learning groups in the classroom.

Slavin (1983) completed an analysis of research articles on cooperative-learning groups to determine which components of the groups were most important to ensure academic improvement. The meta-analysis identified six types of structures: group study with group reward for learning, group study with reward for group product (students worked together and completed a project or presentation for a grade), group study with individual reward, task specialization with group reward for learning (such as Jigsaw II), task specialization with group reward for group product, and task specialization with individual reward.

Slavin (1983) concluded from his analysis of the articles, that the most important aspect of cooperative groups, in order for them to be maximally successful, is individual accountability within the group. In other words, each group member must contribute in a quantifiable way for the group to receive a reward. Unlike group contingencies in which it is difficult to determine the exact contribution of each member (such as the group completion of a project), individual accountability is more likely to inspire each member of the group to work harder for the success of the group. Furthermore, this method appears to encourage high-achieving group members to provide assistance to struggling students, given that all members need to be successful for the group as a whole to be successful. Interestingly, the research does not indicate that group study creates higher student achievement than individual study, as long as there is a group reward with individual accountability (Slavin).

Although most of the research on cooperative-learning groups has been conducted with students in elementary and secondary schools, research on college students has also demonstrated the benefits of using cooperative-learning groups. This research has not identified a link between cooperative-learning groups and overall class participation; however, current research has determined that college students who are learning cooperatively tend to outperform

students who are learning both competitively and independently (Jalilifar, 2010; Johnson, Johnson, & Smith, 1998; Johnson, Johnson, & Smith, 2007).

One recent study (Bowen, 2000) that examined cooperative-learning groups at the college level evaluated the use of cooperative groups in science, engineering, mathematics, and technology courses at the high school and college level. Of the college studies included in the meta-analysis, effect sizes ranged from -0.22 to 0.95. The mean effect size was calculated to be 0.51, indicating that, overall, students experiencing cooperative learning in these courses performed significantly better than those not using cooperative learning. Furthermore, some data were included in several of the analyzed studies on attitudinal variables. According to the research, students who experience cooperative learning had significantly more positive attitudes towards their classes than those who did not.

Not only do academic outcomes and attitudes towards classes improve under cooperative learning at the college level, students report learning skills in addition to mastery of the course content. One study examined student self-reported practical skills after completing a collaborative-learning course versus students who had completed the same course taught in a traditional, lecture-based environment. Students who completed the collaborative course reported significantly more confidence in their ability to apply the knowledge learned in the course to real-life situations, and had more confidence in their ability to communicate orally, resolve conflict situations, be patient with and tolerant of those who have differing opinions, problem solve collaboratively with a group, and recognize inconsistencies and flaws in the thoughts of others and themselves. Even when the researchers controlled for individual student characteristics such as age, race, SAT scores, and high school GPA, students who engaged in

cooperative learning still reported more learning gains than students who completed the course in a traditional format (Terenzini et al., 2001).

Despite studies that have shown the effectiveness of cooperative learning at the college level, many undergraduate classes are still not utilizing the power of cooperative learning. Shimazoe and Aldrich (2010) identified seven aspects of cooperative groups that are frequently met with resistance by students, and may discourage professors from utilizing cooperative groups in their classrooms. At the onset of cooperative groups, students tend to complain that they have had bad experiences with group work in the past, they would rather have a group with their friends instead of being assigned to a group, and that the groups tend to become social and not accomplish any work. During the group work phase, common student complaints include confusion about the assignment, how the group is supposed to function as a unit, and a disproportionate amount of work done by some group members. Finally, at the conclusion of the group work, some students feel as though the group work was just "busy work" and that the grading system, in which every group member receives the same grade, is unfair.

In order to overcome this resistance to group work, Shimazoe and Aldrich (2010) recommended that the professor take a more active role in cooperative-learning process, instead of simply assigning groups and letting students work in their groups with little or no supervision. The authors suggest that professors should ensure that the group task is truly a cooperative task in which students must all contribute in order to complete the group goal, and that students understand why they needed to work together, rather than just telling students that they were required to work as a group. It is also suggested that, if necessary, students be taught social skills before introducing group work so that students feel more comfortable working with new people and have the skills to help the group be productive. Furthermore, students should be given the

opportunity to file complaints against other group members who are either not doing their part in the group or taking over the project and not allowing other members to contribute. Finally, the authors recommended that groups be provided with thorough and immediate feedback on their performance. The rubric for grading should be clear and consistent across groups and across individual members.

Critical Thinking and Participation

A primary goal of college-level instruction is to promote critical thinking skills. Studies have shown that over the course of one semester, critical thinking scores do not tend to change dramatically (Smith, 1977; Tsui, 1999); however, it appears that classes involving little or no class participation may result in a decline of critical thinking skills across the span of a semester (Smith). Furthermore, Smith's analysis revealed that critical thinking skills were positively related to class participation, faculty encouragement of student ideas, and peer interaction.

Research has suggested that students who have strong critical thinking skills tend to demonstrate better classroom performance than students who have weak critical thinking skills. Studies have demonstrated that critical thinking is significantly and positively correlated with exam scores (Williams, Oliver et al., 2003; Williams, Oliver et al., 2004). However, some students with low critical thinking skills still performed well on class exams by virtue of strong note-taking skills (Williams & Stockdale, 2003). The relationship between critical thinking scores and class participation has also been researched. In general, students with high critical thinking skills are less likely to be low responders in class discussion. Furthermore, critical thinking scores and exam scores are moderate to strong predictors of class participation, with students having low critical thinking skills and low exam scores tending to fall into low participation groups (McCleary, 2011).

Exam Performance and Participation

Course examinations are an integral part of many undergraduate college courses. In large college courses, multiple-choice exams are often used to assess class knowledge of relevant concepts without creating excessive grading work for instructors. A variety of student factors appear to contribute to performance on exams, including critical thinking and homework completion (Galyon et al., in press). When asked what factors contribute the most to performance on class exams, undergraduates say their self-reported effort is the most important contributor. On the contrary, correlations have demonstrated a stronger relationship between self-ratings of a student's overall ability and teacher input to exam performance than self-reported effort (Williams & Clark, 2004).

Although few research studies have addressed the relationship between class participation and academic performance, a few studies have shown some predictive potential of class participation to exam scores. Galyon et al. (in press) found a significant positive correlation between participation in the class discussion and performance on some exams. In most cases, class participation alone seemed to be a relatively weak predictor of exam performance. However, a combination of class participation and critical thinking scores seems to be a fairly strong predictor of exam performance.

Framework for the Current Study

The current study builds on previous doctoral studies directed toward improving class participation in a large undergraduate course in educational psychology (Aspiranti, 2010; Foster et al., 2009, Krohn et al., 2010; McCleary et al., 2011). Foster et al. (2009) and Krohn et al. (2010) showed increased and more balanced discussion during class units when participation credit was given on all days of the unit. Aspiranti (2010) expanded on this method by offering

credit on only two days for units during which participation credit was available. In one section, students were informed of the units that would offer participation credit, but were not told which days of the unit counted for credit until the end of the semester. In two other sections of the course, students were not informed until the conclusion of the course as to the units during which participation credit was available or the days that counted during the credit units.

The reason for randomizing selection of credit days and delaying credit in the Aspiranti (2010) study was twofold. First, the researcher hoped to decrease the amount of instructor time required for record-keeping. Second, having unknown credit units was hypothesized to balance class participation across the course, as opposed to this effect being limited to known credit units. Aspiranti (2010) concluded that students participated the most when they were informed of the units in which they would receive credit and the days were chosen randomly at the end of the semester. Including unknown credit units in addition to unknown credit days failed to maximize participation within units but did produce more consistency in participation across units.

McCleary et al. (2011) sought a middle-ground between the Krohn et al. (2010) and Aspiranti (2010) studies. In the McCleary et al. research, students were informed as to which units would yield participation credit, but not the specific days within each unit. At the conclusion of each unit, as opposed to at the conclusion of the course, two days were randomly selected from the unit to count for participation credit. Results from this research were similar to those found in Krohn et al. (2010). Randomized credit for known credit units produced fewer non-participants and more credit-level participants than non-credit units, similar to the Krohn et al. study in which students were given credit every day for participation during credit units.

The current research extends previous research by adding the element of cooperative learning (specifically, cooperative-group rewards for participation) in addition to individual

credit for class participation. Although cooperative learning groups should increase participation within the group discussion, there is minimal evidence that experience in the cooperative groups will increase participation in class-wide discussion. Additionally, to my knowledge, group credit contingencies have not been used to increase class-wide participation by all members of cooperative groups. The model of providing individual credit for participation in the current study closely mirrors that of McCleary et al. (2011). Individual participation credit was available during every unit on randomly selected days, and additional group rewards were available during two, pre-announced units. Group rewards in which each individual team member must contribute for the group to receive a reward appears to be a more effective incentive than individual incentives alone in promoting mastery of academic skills (Law, 2008; Slavin, 1983; Slavin, 2011). As such, I hypothesized that team members would be more likely to participate in class-wide discussion when offered group credit for class participation in addition to being individually accountable for their participation. Furthermore, following McCleary et al. (2011), I analyzed critical thinking as a potential covariate for classroom participation. In addition, I assessed the relationship between exam performance and class participation under different credit contingencies.

Chapter II

Method

Participants

Participants were students enrolled in three sections of a spring semester class in educational psychology (n = 56, n = 55, n = 56 for each section) at a large university in the Southeastern United States. Most of the participants were female (73.1%), and most were sophomores (50.9%) and juniors (23.4%). Students enrolled in the course had a self-reported mean GPA of 3.19.

Setting

Class meetings were held on Monday, Wednesday, and Friday mornings for 50 minutes in a small lecture hall. Each class section had different instructors; however, each instructor had a prior year of training in the course materials and was supervised by the same faculty mentor. Each class section followed the same syllabus and course schedule to allow for maximum similarity among the class sections. To further increase similarity between sections of the course, the instructors practiced teaching in small group sessions with feedback from the supervising faculty member and were provided with lists of questions to prompt discussion among class members.

The course was divided into five units that covered various aspects of human development. The structure of the class and course material were similar across units, with most units lasting seven days depending upon the amount and difficulty of the material for that unit. Two of the more difficult units included an extra day for the instructor to cover and review course material. On the first day of each unit, class members viewed a video that introduced a

relevant topic to be covered in the upcoming unit. Days two through five consisted of class discussions related to the course readings. On discussion days, students were expected to arrive to class having reviewed the readings to be covered that day, as well as having completed a set of questions over the material. Day six of the unit involved a review of relevant research articles, as well as a practice exam and general review of the unit before the unit exam on the seventh day.

Recording Procedures

Recording student comments. Along with required reading materials for the course, students purchased packets of 3x5 record cards on which to record their comments on discussion days (days two through five of each unit). In addition to class comments, this record card confirmed that the student was present for class, displayed a name card during class, and had completed the assigned reading and corresponding questions for that day's class discussion. Students were instructed to briefly record their voluntary comments made in class on the record card immediately after making each comment. Voluntary comments were defined as any comment that the student contributed to the discussion, including questions about course content, comments or ideas about the information being discussed, and responses to any instructor questions. The front of the record card had three spaces for comments to be recorded; however, students were instructed to record every comment they made, using the back of the card if necessary. Record cards were turned into the instructor or teaching assistant at the conclusion of the class period. Receipt of credit for daily participation, as well as for the display of their name card and the completion of homework, provided an incentive for students to turn in the record card each day.

On the fourth day of each unit, inter-rater agreement was assessed to ensure accurate comment recording by the students. Due to high agreement between student self-recordings and

observer ratings reported in previous studies (Krohn et al., 2011), I followed the McCleary et al. (2011) procedures in assessing inter-rater agreement on only one day per unit. One teaching assistant from the course collected all data used in computing inter-rater agreement for all three sections of the course. The observer recorded comments made by each student, and judged the comment as either timely (on topic and accurate) or repetitious (inaccurate or off-topic) based on the feedback given by the instructor. To assist the teaching assistant in completing this reliability check, students displayed name tags at their seats and were required to raise their hands if they wanted to make a comment. Course instructors also called on all students by name to ensure that the rater knew which student was speaking.

Recording teacher behavior. Data on teacher behavior were collected on the fourth day of each unit. Two graduate students, who were also instructors for the educational psychology course, completed observations of the teachers' in-class behavior. Observers judged the types of questions the instructor posed to the class (factual or comprehension), as well as the type of feedback provided to students (positive or negative). Factual questions were defined as questions that could be answered by reading directly from the assigned course material. In contrast, a comprehension question required the student to synthesize material from the readings in comparing, contrasting, analyzing, interpreting, and applying knowledge gained from the course materials.

Feedback given by the instructor to students was coded by the raters as being either positive or negative. Instructor feedback was defined as a statement by the instructor to a specific student regarding the accuracy and relevance of their comment in the discussion. Positive feedback indicated to the student that his or her comment was correct or valuable to the class discussion in that it accurately explained or interpreted the topic being discussed. Conversely,

negative feedback specified that the student's comment was incorrect, only partially correct, off-topic, repetitious or irrelevant to the class discussion. Regardless of the feedback provided to the student by the instructor (positive or negative), their comment could still be recorded to receive participation credit.

Participation Credit

Individual credit. Students received individual credit for class participation in four units of the course. At the conclusion of each unit in which individual participation credit was available, two days from the unit were randomly drawn by one of the students from the class. If a student participated at least once on the day drawn, he/she received three points for the first comment, and two additional points for a second comment. Thus, for each unit (2 through 5), 10 points of individual credit were available to students, for a total of 40 points by the end of the semester. Additional credit was available at the end of the semester if the student participated on every day selected for credit across units. If a student participated at least once on each selected day, the student received 5 additional participation points; if a student participated at least twice on each selected day, the student received 10 more participation points. In summation, students could receive up to 50 points of individual participation credit throughout the course, amounting to eight percent of the overall course grade.

Group bonus credit. Students were assigned to cooperative-learning groups based on their levels of participation during the first unit of the course (during which they received no participation credit). Each group was comprised of initially-high, medium, and low participants. Most groups had five members, but several groups contained six members due to the number of students in the course. The average number of comments across the baseline unit for each group ranged from approximately five to six. Groups and group members were identified by the

instructor at the beginning of the second unit of the course, and students remained in the same cooperative groups for the remainder of the semester.

In two selected units, students were eligible to receive bonus credit based on the participation level of their group. In the context of the course, bonus credit was awarded in addition to regular credit, and there was no grade penalty for not receiving bonus credit. During these selected units, on the two days that were randomly selected for individual participation credit, each member of cooperative groups could earn up to 5 points of bonus credit each day if each member of the group participated at least one time on the selected day. Thus, cooperative-learning groups could receive up to 10 bonus points per unit, totaling 20 points for the semester, representing three percent of the overall regular credit points available during the semester. Table 1 shows the treatment conditions (either individual credit only or both group credit and individual credit) for each section of the course. Each section of the course began with a baseline unit, and subsequently alternated phases between individual-credit-only and individual-plus-group credit units. Section A of the course offered group credit in units three and five, while Sections B and C had group credit in units two and four. See Table 1 for a visual depiction of the treatment sequence in each section.

At the beginning of every class after the first unit (baseline), students were reminded of the credit contingencies for the current unit (individual-credit-only or individual-plus-group credit available). Each group was assigned a number to distinguish groups. During units in which group credit was available, group numbers were placed on the tables in the lecture hall where class was held, with similar group numbers being placed as close together as possible. Students were instructed to sit at a number corresponding to their group number. This seating arrangement allowed students to monitor one another's participation during class and form connections with

other students in their group. Students were not required to do any class activities with their group; rather, class members were placed in these groups solely for the purpose of facilitating the acquisition of bonus credit for class participation.

Credit delivery. Students received credit for individual and group participation based on their self-reported records of participation. Class members were informed of their participation credit through a Blackboard™ class website. Students were taught to access the class website on the first day of the course. All class grades were posted through the Blackboard™ system, including grades for all class exams, attendance, quizzes, practice exams, class participation, and other credit opportunities. In addition, many classroom materials were posted on the Blackboard™ class website, encouraging students to access the website more frequently. Immediately after the conclusion of each unit in which participation credit was available, student participation credit grades were posted. Within the online grading system, separate credit categories were created for "individual credit" and for "group credit," so students could distinguish precisely what types of credit they received. Students were asked frequently by the teaching assistants and instructors in the course to check their posted grades for accuracy.

Unit Exams

At the conclusion of each unit, students completed an exam to test their knowledge on concepts covered in that unit. Each exam consisted of 50 multiple-choice items. Each question was worth one point, such that each exam was worth 50 points. Overall, each individual exam was worth approximately 8.5% of the student's overall grade, and all of the exams combined amounted to about 42% of the students' overall grade in the course. Questions on the exam were taken from all aspects of the course material**,** including the video shown in class on the first day of the unit, concepts covered in instructor notes, ideas introduced in journal articles, and

information in PowerPoint slides. Instructors of each section were familiar with the unit exams and were available to answer relevant questions during the exam administrations. Most of the exam questions were not simple recall questions. Previous analysis of the exam items by Wallace and Williams (2003) showed that only 26% of the unit exam items involve pure recall of facts learned during class, whereas 58% of the items require a more conceptual application of the course material to obtain credit for the item. The remaining 16% of the test items involved a combination of recall and application. The exams have also been analyzed for internal consistency of student responses, with internal consistency of .78 for exam items based on instructor notes and .73 for exam items based on course readings.

Critical Thinking Measures

Students were given the short form (Form S) of the *Watson-Glaser Critical Thinking Appraisal* (Watson & Glaser, 1994) on the second day of class. Students earned five points of course credit for the completion of the appraisal. Form S of the Watson-Glaser includes 40 critical thinking items. The *Watson-Glaser Critical Thinking Appraisal* has been used in previous studies to determine the relationship between critical thinking skills and various aspects of academic achievement, including class participation (McCleary et al., 2011). Factor analysis of the Watson-Glaser has revealed that the appraisal measures a unitary construct. Validity has also been established by identifying positive relationships between the critical thinking score as measured by the Watson-Glaser and GPA, SAT scores, as well as scores on the *Collegiate Assessment of Academic Proficiency Critical Thinking Test* (Hassan & Madhum, 2007).

Several research studies within the educational psychology course have examined the predictive potential of the *Watson-Glaser Critical Thinking Appraisal*. There has been a well-documented relationship between high critical thinking scores and high exam performance

(Williams, Oliver et al., 2003; Williams, Oliver et al., 2004). Although many students with lower critical thinking skills may not perform as well on the exams, some low critical thinkers with strong note-taking skills achieve high exam performance (Williams & Stockdale, 2003).

Chapter III

Results

All data collected throughout the semester were entered into an SPSS database. In addition to the students' self-reported daily comments and observer records of both student and teacher comments, data were also included on demographic variables (e.g. age, gender, and academic standing), Watson-Glaser critical thinking scores, and exam scores, Data were analyzed across all sections and credit units in the course. With the exception of data on inter-rater agreement, all data was analyzed using students' self-recorded participation.

Initial analyses consisted of computing inter-rater agreement for student participation and different types of teacher behavior (questions and feedback). Then, I graphed the percent of students in each section who were non-participants, credit-level participants (one to two comments), frequent participants (three to four comments), and dominant participants (five or more comments) each class day. Because this research used a within-subjects design to examine levels of class participation under different contingencies throughout the course, a three-way mixed ANOVA was performed between sections and across the first and second application of the two different contingencies (individual credit and individual-plus-group credit) (see Figure 1).

The dependent variable in the study was the amount of participation by each individual student. The independent variables were the types of credit (individual credit or the combination of group and individual credit) available for class participation, two applications of each credit condition, and three sections of the course. Teacher behavior was also analyzed to determine whether instructor behavior (amount of factual versus comprehension questions) was consistent across sections and units within sections. Additional data from the students, including their

critical thinking scores and exam performance, were analyzed to determine possible relationships between these variables and the effects of the credit contingencies on class participation.

Inter-rater Agreement between Student and Observer Records of Participation

Inter-rater agreement on student self-recording of comments was assessed on the fourth day of each unit. While students' self-recorded comments made each day, one observer recorded comments made by each student on one day of the unit to check for agreement in comment recording. Agreement between students and observers ranged from 67 to 92 percent (see Table 2). Table 3 shows the means and standard deviations of recorded comments by students in each section on the inter-rater agreement check day for each unit. In general, the means reported by students and the observer indicate that students tended to under-report rather than over-report comments. These results are comparable to the findings of previous studies examining inter-rater agreement between student participation recordings and observer recording of comments (Aspiranti, 2010; Krohn et al., 2011; McCleary, et al., 2011).

Inter-observer Agreement for Instructor Behavior

Data on instructor behavior were collected on the fourth day of the unit. Two independent raters collected data based on the types of questions posed by the instructor (factual versus comprehension), and type of feedback given to students (positive, neutral, or negative feedback) (see Table 4). Table 5 shows percent agreement between the two observers for types of questions asked by the instructor. Percent agreement between raters ranged from 41 to 100%. Figure 2 displays the number of total questions posed by each instructor. Section A's instructor maintained similar numbers of total questions throughout the course, with the exception of Unit 2, in which the instructor posed significantly more questions. The Section B instructor maintained similar amounts of questions across units. Section C's instructor showed an increase

in the number of questions posed following the baseline unit, which was maintained throughout the rest of course. Figure 3, which displays the number of factual questions posed by each instructor, shows a similar pattern to the overall number of questions asked. Section A maintained consistent levels of questions with the exception of Unit 2; Section B maintained consistent levels of questions across the course; and Section C increased the number of questions posed after Unit 1. Conversely, Figure 4 displays the number of comprehension questions asked by each instructor across units. The graph illustrates that instructors in Sections A and B asked similar amounts of comprehension questions across all units, while the number of comprehension questions asked by the instructor in Section C steadily increased during Units 1 through 3 and decreased in Units 4 and 5.

Table 6 shows percent agreement between two observers for type of feedback given to students. Percent agreement ranged from zero to 100%. Overall, reliability was greater for positive feedback (ranging from 74 to 100%) than for negative feedback (ranging from zero to 100%). The poorer agreement between raters for negative feedback was likely due to the small amount of negative instructor feedback given to students, drastically reducing the percent agreement for only small disagreements between raters.

Visual Analysis

Visual analyses of results indicated only moderate increases in participation as a result of the addition of individual participation credit and group contingencies in Sections A and B of the course. Figure 5 shows that the mean percent of daily non-participants in Sections A and B to be lower than baseline in all subsequent units. Furthermore, although the drop was modest, the average number of non-participants was lowest in units in which group credit was available in addition to individual credit. This effect can be examined further by looking at the percentage of

credit-level participants in each section of the course (i.e., students participating one or two times per class period). Again, while the effect is mild, mean percentages of credit-level participants in units providing credit for participation was consistently higher than baseline rates in Sections A and B (see Figure 6). Furthermore, the mean percentage of credit-level participants was slightly higher during individual-plus-group credit phases than during individual-only credit phases. Section C did not show the same pattern as Sections A and B. As seen in Figure 5 and Figure 6, Section C did not consistently have fewer non-participants or consistently more credit-level participants in Units 2 through 5 than in baseline.

Figures 7 and 8 show the percent of frequent participants (three to four comments per day) and dominant participants (five or more comments per day) in each section of the course. In general, all sections show a steady decrease of frequent participants across the semester. There does not appear to be a relationship between the different credit contingencies offered in the course and the percent of individuals commenting three or four times per class period. Similarly, Figure 8 shows no obvious relationship between credit contingencies in the course and the percentage of dominant participants. In both Sections A and C, levels of dominant participants in the course never exceeded baseline levels. More specifically, in Section A, none of the data points in Units 2 through 5 reached the highest percent of dominant participants during baseline. In Section C, most participation days remained below the highest percent of baseline participation, with one treatment day having the same percent of dominant participants as during the first unit. While most of the data points in Section B indicate that the credit contingencies kept the percent of dominant participants in the course at low levels, there were three class days during Units 2 and 3 in which the percent of dominant participants in the class slightly exceeded any of the baseline levels of dominant participants. While it is difficult to ascertain exactly why

more students participated at dominant levels on these days, perhaps some students found the discussion more engaging and became more involved in the class discussion on these days. Furthermore, two of the days with higher levels of dominant participants were inter-rater check days during which three graduate students observed the class. The presence of the observers on these days may have changed the dynamics of the class and created a slight increase in percentage of dominant participants. Overall, visual analysis shows that dominant class participation was virtually non-existent under all conditions.

Comparison of Treatment Data Points to Baseline Median Levels

Traditional methods of visual data analysis involve computing the percentage of non-overlapping data points between baseline and treatment phases of the research study. However, because the baseline, treatment, and removal phases of data consisted of high amounts of variability, tabulating the percentage of data points in treatment conditions that exceeded the median (PEM) in baseline may be a more appropriate way for analyzing the data. Previous studies comparing quantitative data analysis methods in single-subject designs show that determination of treatment effects is better captured by the PEM than by the more traditional method of non-overlapping data points (Ma, 2006; Wolery, Busick et al., 2008).

I calculated the median percentages of non-participants during the baseline unit for each section. Across the first (baseline) unit in Section A, the median percentage of non-participants was 36.8. In Sections B and C, the median percentages of non-participants in the baseline phase were 35.25 and 17.7 percent, respectively. These median percentages are denoted by a dotted line in Figure 5. In Section A, 100% of the data points for non-participants during group-plus-individual credit phases were below the median level of baseline (indicating less non-participants), while 88% of the data points in the individual-credit only phases of the study were

below the median level. Similarly, in Section B, 100% of the data points during group-plus-individual credit phases were below the median baseline level, while only 63% of the data points were below the median baseline level when only individual credit was offered. These results suggest a slight advantage of using group credit in addition to individual credit to decrease the percentage of non-participants. Section C of the course showed less promising results, with only 12.5% of the non-participant points occurring below median baseline levels during individual-plus-group credit units and only 12.5% of points occurring below median baseline levels during individual-credit-only phases of the study. Individual credit appears to have had a similar effect to individual-plus-group credit in Section C, with both effects similar to those of baseline conditions.

In order to further assess the effectiveness of combined group and individual credit over individual credit only, median percentage of non-participants in the individual-credit-only phases was compared to the levels of participation in the group-plus-individual credit phases. In Section A, the median level of non-participants across both individual-credit-only phases was 28.15 percent. In Section B, the median level of non-participants across both individual-credit-only phases was 27.55 percent. Lastly, in Section C, the median level of non-participants across the individual-credit-only phases was 21.95 percent. Section A showed positive results with 100% of the data points in the group-plus-individual credit phases being below the individual-credit only median level, indicating consistently fewer non-participants in the combined credit phases. Section B showed slightly less positive results, with 87.5% (seven out of eight data points) of the points during the group-plus-individual credit phases falling below the median level of non-participants during the individual-only-credit phases. Section C had split results, with only 50%

of the data points in the group-plus-individual credit phases falling below median level of non-participants in individual-credit-only phases.

Median levels of participation during the individual-credit phases were also calculated for credit-level participants (one to two comments daily). In these analyses, unlike the previous PEM analyses involving non-participants, percentages of points above the median level of baseline participation indicated positive results. Section A had a median level of 51.8% of the class falling at credit-level participation in the individual-credit-only phases of the study. Similar to results of the previous PEM analyses in Section A, 100% of the data points for percentage of credit-level participants in the individual-plus-group contingency phases were above this median level of participation, signifying consistently higher levels of credit-level participation in the individual-plus-group credit units in this section. In Section B, the median level of credit-level participants during the individual-only-credit phases was 48.2% of the class. In the individual-plus-group credit phases, 100% of the data points exceeded this median level of performance. Finally, in Section C, the median percentage of credit-level participants in individual-only credit units was 53.6%. Examination of the individual-plus-group credit phases of the study shows that 37.5% of the data points in these units exceeded the median level of the individual-credit-only phases.

Class-Wide Statistical Analysis of Participation

In order to determine the effects of the two different interventions across time and sections, I used a three-way mixed analysis of variance design. This ANOVA assessed the impact of the two different within-subject variables across time (at the first and second implementations of each type of credit contingency) and between the three sections of the course (see Figure 1). No significant interaction effects between any of the variables were found for

total participation (see Table 8). No significant main effect was obtained for section, $F(2, 154) = .376$, $p = .688$, but there were significant main effects for contingency type, $F(1, 154) = 6.117$, $p = .014$, and application time, $F(1, 154) = 31.448$, $p < .001$. The mean level of participation was higher during the group-plus-individual credit contingency than during the individual-only-credit contingency. With respect to the main effect for treatment applications, the first application of the credit contingencies produced significantly more participation than the second application of the credit contingencies.

Figure 9 displays the effects of the group and individual credit contingencies across the first and second applications of credit for all three sections of the course. While the effects are significant, the graph shows relatively small differences between amount of participation during the individual-only-credit contingency and the individual-plus-group credit contingency. When the less-responsive Section C is dropped from the analysis, the main effect of contingency slightly increases in significance, $F(1, 154) = 23.074$, $p = .011$, indicating greater mean differences in participation between individual-plus-group credit and individual-only credit phases in the A and B Sections. Examination of the means show that individual-plus-group credit continues to promote greater participation than individual-only credit in these two sections.

Cross Unit Participation of Initially-Low Participants

Although the intervention was used with the entire class, students initially reluctant to participate were the primary targets of the intervention. In order to examine the effects of offering individual participation credit versus individual-plus-group participation credit on these initially-low participants, the classes were separated into quartiles based on the total number of comments made during the baseline phase of data collection. For Sections A and B of the course, the bottom quartile of students represented students who participated an average of .25 times per

day during baseline or less. In Section C, the bottom quartile of students represented students who participated an average of once per day during baseline or less. Analysis of initially-low participants included only these students who fell within the bottom quartile of participation for their section.

In order to determine the effects of the two different interventions across time and across sections, the same three-way mixed analysis of variance, as described above, was repeated with the initially-low participants. This ANOVA, just as with the ANOVA conducted on the entire class, assessed the impact of the two different interventions (individual credit and individual-plus-group credit contingencies) across time (at the first and second implementations of each type of credit contingency) and between the three sections of the course (see Table 9). One significant interaction was found between all three variables (section by time of application by credit contingency), $F(1, 41) = 5.389$, $p = .008$. Analysis of simple effects shows that Sections A and B showed similar patterns of contingency effects as with the whole-class participation (see Figure 10). However, the contingency effects were much more substantial for the initially low participants in Sections A and B than for class wide participation across sections (as portrayed in Figure 9). Individual-plus-group credit produced consistently higher rates of participation than individual credit alone and the first application of credit produced greater participation than the second application of credit. Section C showed a vastly different pattern with little difference in participation levels between times of application or types of credit offered (see Figure 11). Across both the first and second application of credit contingencies, no significant mean differences were found between the level of participation during the individual-only-credit phases and the individual-plus-group credit phases. This interaction effect supports previous analyses indicating that Section C produced a limited response to both interventions.

Due to the significant interaction between sections, time of application, and credit contingencies, within-subjects analyses were performed to determine which specific units and credit contingencies created the greatest participation among the initially-low participants. An ANOVA was performed with the Section A initially-low participants to see if there was an overall significant difference between the two types of credit contingencies in this section, irrespective of application time. The ANOVA showed a significant effect for contingency type (individual-only or group-plus-individual credit), $F(1, 11) = 6.681$, $p = .024$, with individual-plus-group credit producing significantly higher rates of participation than individual-only credit.

Further analysis with Section A initially-low participants indicated that there were significant differences in amount of participation across units, $F(4,11) = 14.505$, $p < .01$. Bonferroni post-hoc analyses were performed to identify significant differences between specific credit units. During Unit 2, the first unit of individual credit, the initially-low responders participated significantly more per day than during baseline, $p = .02$. Under the combined group-plus-individual credit during Unit 3, initially-low responders still participated more than baseline, $p < .01$. Unit 4, in which only individual credit was offered, showed a decrease in participation, with participation not significantly different from baseline participation, $p = .12$. Also, participation in Unit 4 was significantly lower than in Unit 3 when group credit was available, $p < .01$. When group credit was offered again in addition to individual credit during Unit 5, levels of participation was again higher than baseline, $p < .01$. However, amount of participation during Unit 5 was not significantly different from participation during any of the other credit units. Overall, these results within Section A illustrate that providing group credit plus individual credit always produced higher participation than baseline, whereas offering only

individual credit for participation only yielded higher rates of participation than baseline during the first individual-credit-only unit.

This same procedure was used to analyze participation patterns of initially-low participants within Section B. An ANOVA was performed with the Section B initially-low participants to see if there was an overall significant difference between the two credit contingencies in this section. The ANOVA showed a significant effect for contingency type (individual-only or group-plus-individual credit), $F(1, 15) = 7.070$, $p = .018$, with individual-plus-group credit producing significantly higher rates of participation than individual-only credit.

There was a significant difference between amount of participation based on unit, $F(4, 15) = 12.95$, $p < .01$. Bonferroni post-hoc analyses were performed to identify differences across specific units. During Unit 2, when students were offered both individual and group credit, the initially-low participants averaged significantly more comments per day than during baseline, $p < .01$. During Unit 3, when students were only offered only individual credit, reluctant speakers still participated significantly more than baseline, $p = .02$. When group credit was added back in addition to individual credit during Unit 4, these students continued to participate more than baseline, $p < .01$. In Unit 5, when only individual credit was offered, participation dropped across the initially-low participants, with participation among these students not significantly different than during baseline, $p = .27$. This analysis demonstrates similar effects of the contingencies in Section B as in Section A. With regards to initially-low participants, the individual-only phases of the study produced higher rates of participation only in the first unit when individual credit was offered, while the individual-plus-group phases consistently produced greater participation than baseline.

Section C was analyzed by using the same procedure with the bottom quartile of Section C students, who were participating an average of one time or less per day during baseline. An ANOVA was performed with the Section C initially-low participants to see if there was an overall significant difference between the two credit contingencies in this section. The ANOVA did not show a significant effect of contingency type (individual-only or group-plus-individual credit), $F(1, 15) = 0.60$, $p = .810$. Unlike the previous two sections, there was no significant difference between amount of participation based on unit, $F(4, 15) = 2.277$, $p = .07$. It should be noted, however, that the bottom quartile of Section C participated more frequently during baseline than the bottom quartile of the previous two sections. In fact, 30% of the students identified as being in the bottom quartile (or 9% of the overall class) in Section C participated an average of once a day during baseline phase. To determine if the intervention was effective on students participating less than once a day, the repeated measured within-subjects analysis was repeated using only students who participated, on average, less than once per day. With this adjustment, the analysis showed a significant difference in average amounts of participation based on unit, $F(4,15) = 2.604$, $p = .049$. However, Bonferroni post-hoc analyses did not reveal any significant differences between units.

Critical Thinking

I examined critical thinking as a potential covariate of classroom participation. Mean critical thinking scores were first calculated for each section of the course. Section A and B had similar mean critical thinking percentile scores, with mean scores falling at the 26th and 28th percentiles, respectively. Section C had higher mean critical thinking scores, with a mean class percentile of 37. The raw critical thinking scores were closer together in means than the percentile scores (Section A = 26.63, Section B = 26.81, Section C = 27.60). To determine if the

mean critical thinking scores in the three sections were significantly different from one another, a one-way ANOVA test was performed on the critical thinking raw scores. The ANOVA analysis revealed no significant difference between sections, $F(2, 159) = .381$, $p = .68$.

Some previous research has indicated that higher critical thinking skills are related to higher rates of class participation. McCleary (2011) found that higher levels of critical thinking are moderate predictors of greater class participation. Thus, critical thinking raw scores in the current study were correlated with baseline levels of participation to examine the relationship of critical thinking to class participation in the absence of credit for participation (see Table 10). In Section A, critical thinking was not significantly correlated with baseline levels of participation ($r = 0.21$, $p = 0.16$). Section B showed critical thinking to be moderately correlated with baseline levels of participation ($r = 0.43$, $p = .003$). Interestingly, in Section C, critical thinking was negatively correlated with baseline levels of participation. However, this relationship did not reach statistical significance ($r = -0.19$, $p = 0.08$). Table 10 also shows correlations between critical thinking and levels of participation across the entire semester.

Not finding significant differences in critical thinking across sections combined with obtaining a highly unusual correlation between critical thinking and baseline participation in Section C led me to forego the inclusion of critical thinking as a covariate in any further data analysis. The negative correlation between critical thinking and class participation (obtained in Section C) has not been reported in any other study (McCleary et al, 2011; Smith, 1977). Thus, Section C appears to represent an outlier with respect to the relationship between critical thinking and participation. Consequently, including critical thinking as a covariate would more likely confuse rather than clarify effects of the treatment conditions on participation.

Data were also analyzed in all sections regarding the correlational relationship between critical thinking and exam scores (Table 11). Correlations between critical thinking skills and exam performance were positive and significant across most exams. It is important to note that while the correlations are all positive, the correlations tend to be weaker and less significant in Section C than in the other two sections of the course.

Participation and Achievement Patterns within Sectional Groups

Section A. Participation groups were created such that each group reflected a diversity of baseline participation levels but similar mean levels of baseline participation. Other course factors such as class performance were not factored into the creation of participation groups. For the purpose of examining the similarity of groups, mean achievement levels on the Unit 1 exam were also analyzed across groups within sections. Section A of the course consisted of 11 participation groups. The mean group performance on the baseline exam (Unit 1) ranged from 36.4 to 43.8 (see Table 12). A one-way ANOVA showed no significant differences among the mean baseline exam scores for Section A groups. A repeated measures analysis did not show any significant difference between unit exam scores for any group in Section A. Table 12 also designates the amount of credit that each group received for group participation across the entire semester. In Section A, there was no apparent relationship between initial level of class achievement and total amount of group credit obtained during the semester.

Analyses were also conducted in Section A to see if students who received credit for class participation tended to perform better on exams than those who did not receive credit for participation. In Units 2 and 4 (individual-only-credit units), students who received participation credit did not perform significantly better on exams than students who did not receive credit. Within the group credit units, 82% of groups received bonus credit during Unit 3, and 64% of

groups received bonus credit in Unit 5. Within both of these individual-plus-group credit units, individuals who received group participation credit did not perform significantly better on exams than students who did not receive credit.

Section B. Section B of the course consisted of 11 participation groups. The mean group performance on the baseline exam (Unit 1) ranged from 38.2 to 44.8 (see Table 13). A one-way ANOVA showed no significant differences among the mean baseline exam scores for Section B groups, $F(10, 44) = 1.00$, $p = .457$. A repeated measures analysis did not show any significant difference among exam scores across units for any group. Table 13 also denotes the amount of credit that each group received for group participation across the entire semester. Again, there does not appear to be a clear pattern between amount of group participation credit received and baseline exam achievement levels.

Analyses were conducted in Section B to see if students who received credit for class participation tended to perform better on exams than those who did not receive credit for participation. In Units 3 and 5 (individual-only-credit units), students who received participation credit did not perform significantly better on exams than students who did not receive credit. Within the group credit units, 27% of the groups received bonus credit during Unit 2, and 45% of groups received bonus credit in Unit 4. Within both of these individual-plus-group credit units, individuals who received group participation credit did not perform significantly better on exams than students who did not receive credit.

Section C. Finally, Section C of the course consisted of 11 participation groups. The mean group performance on the baseline exam (Unit 1) ranged from 33.80 to 40.00 (see Table 14). A one-way ANOVA showed no significant differences among the baseline exam-score means for Section C groups, $F(10, 44) = 0.64$, $p = .771$. Table 14 also shows the amount of credit

that each group received for group participation across the entire semester. A significant relationship did emerge between baseline and group-bonus-credit achievement levels. The two groups (Groups 28 and 31) who received any group credit for participation also received the lowest mean scores on the baseline exam. However, as indicated by the previous analysis, their performance on the baseline exam was not significantly lower than the performance of the other groups in Section C.

A repeated measures analysis was performed to examine differences in exam scores within groups across units. In Section C, several groups showed significant differences between exam scores across different units. Group 23 showed significant differences between exam scores based on unit, $F(1, 4) = 13.470$, $p = .021$. The Unit 2 exam score was significantly different from that of the Unit 5 exam score. Group 27 also showed significant differences between exam scores based on unit, $F(1, 4) = 8.897$, $p = .041$. Bonferroni post-hoc analyses revealed a significant difference between the Unit 2 exam score and the Unit 3 exam score in this group.

The two groups (Groups 23 and 27) that had significant differences between exam scores across two units did not receive any credit for group participation. In both Groups 23 and 27, the significant difference between exam scores fell between an individual-plus-group credit phase and an individual-only-credit phase of the course. To analyze if level of participation within the group was significantly different between the two units in which exam performance differed, I performed a repeated measures analysis. Group 23 did not show a significant difference between the amount of participation in Unit 2 and Unit 5. Likewise, Group 27 did not show a significant difference between the amount of participation in Unit 2 and Unit 3. To further ensure the non-effect of participation on exam performance in these two groups, both groups were compared to

a group that had higher performance on the Unit 2 exam, Group 30. There were no significant differences between the amount of participation in the higher performing group (Group 30) and the lower performing groups (Groups 23 and 27). Thus, while the groups differed with regard to their performance on the Unit 2 exam, participation across all groups was similar.

Additional analyses were completed to see if students who received credit for class participation tended to perform better on exams than those who did not receive credit for participation. In Units 3 and 5 (individual-only-credit units), students who received participation credit did not perform significantly better on exams than students who did not receive credit. Within the group-credit units, none of the groups received bonus credit during Unit 2, and 18% of groups received bonus credit in Unit 4. In Unit 4, individuals who received group participation credit performed significantly better on exams than students who did not receive credit, $F(1, 52) = 4.860$, $p = .032$. On average, students who received group bonus credit for participation in Unit 4 performed 5 points higher on the exam than students who did not receive group bonus credit.

Summary. Overall, group analyses appear to indicate that most participation groups were highly similar in terms of unit exam performance within each section of the course. Furthermore, a clear connection between exam performance and the achievement of group bonus credit did not emerge from the analyses. Only in Unit 4 of Section C did a relationship emerge between these two variables, with students who received group bonus credit performing better on the unit exam than those students who did not receive group bonus credit.

Chapter IV

Discussion

The purpose of this research study was to examine the effects of cooperative-learning contingencies on participation in class discussion. Undergraduates in the study were students in three sections of a course in Human Development in Educational Psychology. Students were given individual credit for participating in class during all units, with the exception of baseline. During two designated units, students were also given credit based on all members of an assigned cooperative-learning group participating in class on randomly selected days. Composition of the cooperative-learning group was based on baseline levels of class participation and consisted of five or six members per group. Each cooperative group was balanced for baseline participation across students such that each group contained members who had been identified as initially-high, initially-medium, and initially-low participants during the baseline phase. This study sought to examine if a combination of credit contingencies (individual with group credit) was more effective than individual credit only for participation as outlined by McCleary et al. (2011).

Effect of Credit Contingencies on Class Participation

McCleary et al. (2011) offered individual credit for participation during two selected units of each section of the course. Students were informed at the beginning of each unit whether the current unit yielded credit or was a non-credit unit. However, although the unit lasted four participation days, McCleary et al. only offered credit on two days of the unit, which were randomly selected at the conclusion of the unit. Analysis of the results indicated that providing random individual credit for participation decreased percentages of non-participants and increased percentages of credit-level participants.

McCleary et al. (2011) found that random, individual credit increased credit-level participants from approximately 20 to 30 percent of the class to 40 to 60 percent of the class. Similarly, random individual credit decreased the number of non-participants from about 40 to 60 percent of the class to only 20 to 30 percent of the overall class. In the current study, similar results were found within two of the three sections of the course with respect to individual-credit-only units. Sections A and B showed parallel decreases in percentage of non-participants and increases in percentage of credit-level participants in the individual-credit-only phases as did the McCleary et al. research.

An important aspect of this study was to compare individual-credit-only phases (identical to those in the McCleary et al. study) to treatment phases in which individual-plus-group bonus credit was available for participation. Sections A and B of the course demonstrated a small treatment effect for the addition of group credit. In these two sections during group-plus-individual credit units, non-participants accounted for approximately 10 to 20% of the overall class, while credit-level participants account for approximately 60 to 70 % of the class. An analysis of the data using the percentage of points exceeding the baseline median demonstrated higher participation in combined credit phases than during individual-only-credit phases. In Sections A and B, 100% of the points in the group-plus-individual credit phases exceeded the credit-level participation median in individual-credit-only phases. Likewise, 87.5 to 100% of the points in the individual-plus-group credit phases were less than the median percent of non-participants in the individual-credit-only phases.

McCleary et al. (2011) did not show any effect of the credit contingencies on percentage of frequent participants. A similar non-effect emerged in the current study. There were high numbers of overlapping data points between all phases of the study (baseline, individual credit

only, and combined individual and group credit), and the mean levels of frequent participants did not show any consistent relationship with the treatment phases in any section of the course. McCleary et al. found that the percentage of dominant participants in each section maintained similar, near-zero levels throughout all phases of the course. While the current study had slightly higher levels of dominant participants in the class during the baseline unit, data analysis showed similar results in other units, with all sections of the course showing very low percentages of dominant participants for the duration of the course.

Overall, it appears that offering credit for individual and group participation has the greatest effects on the levels of non-participants and credit-level participants. Similar to analyses of Aspiranti (2010), Krohn et al. (2010) and McCleary et al. (2011), this study revealed that the credit contingencies had little effect on the percentage of frequent participants within the various phases of the class. Students already engaging in the discussion and exceeding the credit requirements, but not dominating the discussion by offering excessive comments, may represent students who are highly involved in the course content and engage in participation for their own learning, as opposed to primarily for the course credit.

There was also a minimal effect of the credit contingencies on the dominant (five or more comments per class) participants. I hypothesized that the level of dominant participants would decrease with the addition of course credit for participation, because the self-recording procedure would raise the student's awareness of their high amount of participation. One of the reasons that the data shows little change in the amount of dominant participants throughout the course is that the level of dominant participants during the baseline phase of the study was at near-zero levels, not allowing for a significant decrease in dominant participants during the treatment phases of the study.

To further analyze the precise changes in participation levels of individual students, a three-way mixed analysis of variance was performed between the three sections of the course and across the two applications of two types of participation credit (individual-only versus individual-plus-group credit). The analysis yielded no significant interaction effects, no significant main effect for sections, but significant main effects for both contingencies and treatment applications. The ANOVA indicated that the first application of credit had a stronger effect on participation than the second application and that group-plus-individual credit contingencies created higher class participation than individual credit alone (see Figure 9). A potential reason for the greater initial effect of the credit contingencies could be greater motivation to obtain all available course credit at the onset of the class. If students were performing well towards the end of the course, they might have become less motivated to obtain all available course credit, especially if they could obtain their desired grade without participation credit, which was worth only a small percent of their overall grade. Furthermore, the greater effect of credit at the first application may represent a novelty effect. Thus, students improved their performance because credit for participation was new, but this newness did not last throughout the semester.

Effect of Credit Contingencies on Initially Low Participants

Given that the primary focus of the contingencies was to increase participation among initially low participants, each section of the course was divided into quartiles such that the lowest participation quartile of the class in baseline could be tracked across subsequent units. An ANOVA indicated that within Sections A and B, the first application of the credit contingencies was more effective than the second application of the credit contingencies and that group-plus-individual-credit phases produced higher participation than individual-only-credit phases.

Further analysis showed that, in both Sections A and B, both applications of group-plus-individual credit for participation yielded significantly higher rates of participation than baseline phases of class participation in both applications. Furthermore, in both sections, receiving only individual credit for participation was significantly better than baseline in only one of the two units in which it was offered. To further assess the treatment effect of adding cooperative-group credit, I found that percentage of non-participants was less than the median percent of non-participants during the baseline phase of the study. All of the data points collected in the group-plus-individual credit phases of Sections A and B showed fewer non-participants than the median of the baseline phase.

While the effect of providing group-plus-individual credit was not as large as expected, analyses demonstrated that there was a positive effect of providing group credit in addition to individual credit for participation. Within Sections A and B of the course, the individual-plus-group credit contingency produced participation levels that were higher than baseline during both applications of credit, while the individual-credit-only contingency created higher levels of participation in only the first application of credit. Although individual-plus-group credit was always significantly higher than baseline levels of participation, the individual-plus-group credit did not consistently create significantly higher levels of participation than individual-credit alone. Section A was the only section of the course that demonstrated significantly higher rates of participation during individual-plus-group credit compared to individual-credit alone, but this difference only occurred between Units 3 and 4.

Effect of Instructor Behavior of Class Participation

In order to ensure similarity between sections, the questioning-style of each instructor was examined on the fourth day of each unit. During the first two units of the semester, the

instructors questioning styles were the most disparate. In the first unit, the Section C instructor asked fewer total questions than in the other two sections and, in particular, asked far fewer factual questions than the other two sections. In Unit 2, the Section A instructor stood out as asking more total questions and more factual questions than the instructors in the other two sections of the class. During the last three units of the course, the instructors were much more similar in terms of their questioning of the class.

Similar to the conclusions of previous studies on class participation in this educational psychology course (Aspiranti, 2010; Krohn et al, 2010; McCleary et al., 2011), findings of the current study appear to indicate that the instructor's questioning and feedback style had little effect on the students' opportunities to participate. First, as demonstrated by the aforementioned studies and current research, the vast majority of feedback given to students by all instructors was positive. Furthermore, the number of questions asked by the instructor did not appear to directly influence the opportunities for class members to participate. Even in the section that asked the fewest total questions (Section C during Unit 1), students still participated at high levels. Furthermore, there is not necessarily a one-to-one correspondence between the number of questions asked and the number of opportunities for class members to participate in the discussion. For example, if a student gives only a partially correct answer to a factual question, other students may volunteer additional information or clarification to amplify the correctness of the response to the question. In addition, comprehension questions posed by the instructors could have offered many opportunities to participate, by provoking higher-order thinking, and allowing students to discuss their individual analyses of important educational issues.

Relationship of Critical Thinking to Participation

Students completed the *Watson-Glaser Critical Thinking Appraisal* at the start of the semester. Previous studies have indicated that there is a relationship between critical thinking and class participation. McCleary et al. (2011) found that students with higher critical thinking skills tended to be higher participants than students with lower critical thinking. In order to determine the relationship between critical thinking scores and participation in the different sections, I ran Pearson product correlations between critical thinking and participation. The analysis in Sections A and B showed positive correlations between critical thinking scores and baseline levels of class participation. The relationship between these two measures was small in Section A, while the relationship in Section B was moderate in size. In contrast, Section C revealed a non-significant, negative correlation between these two measures. Thus, in Section C, students who were participating more frequently in baseline had lower critical thinking skills than those who were more reticent.

Relationship of Achievement to Participation

A relationship has been documented between class achievement and class participation. McCleary et al. (2011) found that undergraduates who have higher rates of participation tend to have higher rates of achievement on exams. Students who are participating may be paying better attention to class discussion and gain a more complete understanding of the course material than those who are not engaged in the discussion. One of the primary reasons of promoting higher rates of class participation is to subsequently increase class performance as a whole. In the current study, the level of class achievement tended to remain stable across the semester, despite the increase of class participation during units with participation-credit contingencies. Additionally, in Sections A and B, students who received credit for participation (across all types

of credit contingencies) did not perform significantly better on the exam than those students who did not receive credit. Section C did demonstrate a relationship between achievement of group bonus credit and higher performance on the exam during Unit 4, with students who received group bonus credit attaining higher exam scores.

One of the likely reasons for the limited relationship between increased participation and increased class performance was the impact of credit for participation. McCleary et al. (2011) classified students into high and low participation groups based on their baseline exam scores. Participation of high- and low-performing students in the class became more balanced across students as a result of credit for participation. When not provided credit, students who are more invested in the course (more interest and better study habits) may be more inclined to participate; whereas when credit is offered for participation, more students may be inclined to participate in the class discussion simply to receive the participation credit. Thus, high- and low-achievers may participate at similar levels when credit is offered for participation.

Benefits and Drawbacks of Group Credit

One of the main research questions in this study was to examine the effectiveness of the group-plus-individual credit contingency versus individual-credit-only contingencies in increasing participation. Previous research on the individual-credit contingencies has shown moderate increases as a result of individual credit for participation (Aspiranti, 2010; Krohn et al., 2010; McCleary et al., 2011). I hypothesized that the addition of group credit would amplify the effects of the credit contingencies and create greater balance in participation across students in the cooperative groups. Although the results indicated some advantages of using group-plus-individual credit over individual credit alone, the benefits were not as pronounced and consistent across sections as had been expected.

Despite the limited effectiveness of the group contingencies, several positive outcomes of the group contingencies emerged. First, anecdotal reports from the teachers indicated that the group contingencies clearly encouraged participation in some otherwise reticent students. For example, the Section A instructor reported that, within some groups, students would verbally encourage each other to participate in class. Furthermore, some groups talked at the beginning of class to ensure that all members of the group were prepared to participate and had a planned comment to make. It can be argued that having a "planned comment" at the beginning of class could be a drawback, such that students do not learn how to naturally participate in class. However, for some particularly reticent students, having a planned comment could be comforting. In addition, once these students learn to offer a planned comment, they may begin to volunteer comments more spontaneously.

The individual-plus-group contingencies were ineffective with some students. Some students remained non-participants despite the bonus credit available if all members of the group engaged in class discussion. Even in Section A, where the group contingencies appeared to have the strongest effect, no less than 7% of the class remained non-participants on any given day. Students in groups who had members who refused to participate, despite credit contingencies, may have become highly frustrated with those members and the contingencies. In order to prevent students from feeling as though their group arrangement was unfair, all group credit was offered as bonus credit. Therefore, if one member of the group continually refused to participate, there was no adverse effect on the grades of other group members. However, it is possible that group contingencies could have been more effective if the group credit was offered as regular credit that directly impacted the grades of all the members of the group.

A second drawback to the use of cooperative learning groups is the lack of supportive relationships developing between group members. The cooperative-learning groups were formed solely for the purpose of providing bonus credit for participation of all group members. While groups were required to sit together during units that yielded group credit, these students were not required to work together during or outside of class. Some groups, in spite of this lack of interaction, did form relationships as evidenced by the anecdotes from instructors stating that group members would discuss their participation before the start of class. Because a sense of interdependence between members is an integral part of the effectiveness of cooperative groups (Slavin, 1983), providing more opportunities for the group to work together both in and outside of class could have improved group dynamics.

Furthermore, in order to make group-credit contingencies more effective, it may be important to provide instruction to students on how to best work with each other within the context of the cooperative groups. As suggested by Shimazoe and Aldrich (2010), many students who are initially reluctant to participate and engage with cooperative groups may simply need instruction in cooperative group work. For example, the instructors can explain, more specifically, the purpose of the cooperative group, and could provide guidance on how the group should deal with members who are uninvolved in the group work or continually refuse to participate (Shimazoe & Aldrich). In addition, some students may benefit from having the opportunity to practice participation with their group members outside of class. Requiring the groups to engage in cooperative-group discussion sessions could provide more reticent students the opportunity to practice talking about course topics in a smaller, more informal environmental setting. Perhaps this practice, and increased familiarity with the course content as a result of the discussion, could provide low participants with confidence needed to participate in class.

Unresponsiveness of Section C

Despite the moderate success of the credit contingencies in Sections A and B, Section C remained unresponsive to the addition of credit for class participation. To explain this unresponsiveness to the credit contingencies, I examined several possible contributors to this pattern. First, differences in instructor behavior were analyzed. During baseline phases of the study, the Section C instructor asked significantly fewer total questions than the other two instructors (12 questions compared to 26 and 30 questions asked by the instructors of Sections A and B). Furthermore, during the baseline phase of the study, the instructor in Section C asked fewer factual-type questions than the other two instructors. In fact, during the baseline inter-rater day, Section C's instructor was recorded as asking only one factual question and 11 comprehension questions. However, as the course progressed, Section C instructor's balance between factual and comprehension became more similar to that of the other instructors. It follows that because Section C had fewer questions posed to the class during the baseline unit, that more students would not have a chance to participate in class. However, this pattern did not follow. During the inter-rater day (day four) of baseline, students in Section C participated more, on average, than students in either of the other two sections (see Table 7). This pattern of higher rates of participation despite fewer questions posed by the instructor implies that students in Section C may have been able to participate more without constant support and questioning from the instructor. Furthermore, inasmuch as the instructor in Section C asked significantly more comprehension questions (which may be more likely to foster an in-depth discussion) than factual questions (which typically have one right answer and may offer only one student the chance to participate), a larger number of students may have had opportunities to respond to each question.

Second, baseline levels of participation are a likely reason for the lack of increase in participation in Section C. The average levels of participation during the baseline phase for Section C were higher than either of the other two sections (see Table 7). Furthermore, the lowest 25% of participants were commenting an average of one time or less a day in Section C. Figure 5 shows that Sections A and B had far more non-participants during the baseline unit than Section C. Sections A and B had similar levels of mean non-participants (37% and 35% respectively), while Section C had a mean of 19% non-participants during the baseline. Thus, adding credit for participation did not provide a great incentive to participate more frequently when most students in the course were already participating at credit-producing levels during baseline without credit.

Next, critical thinking scores were examined as a possible contributor to Section C's higher baseline participation and lack of increase in participation across contingencies. Initial analysis of the critical thinking skills of the students within the three sections showed that critical thinking scores were similar across all sections. Furthermore, a negative correlation was found between critical thinking scores and baseline participation levels and near-zero correlations were found between critical thinking and participation levels during all other units, indicating that critical thinking is not a sufficient explanation for Section C's lack of responsiveness to the participation contingencies.

Another hypothesis for the lack of increased participation during the credit contingency units in Section C is the lack of motivation to obtain participation credit. The individual participation credit available to students represented only eight percent of the overall grade. Thus, assuming students were performing well in all other areas of the course, students could still receive an A or a B grade without having to participate in class. Furthermore, the group

participation bonus credit was worth just three percent of the overall grade. This small amount of credit may not have been significant to many students in the course, particularly those who were maintaining high class averages through attendance, homework, tests, and quizzes. Furthermore, the low motivation of these students to obtain course credit might be explained by the time of the class. Section C was the only researched section that took place entirely in the afternoon. Sections A and B began at 10:10 am and 11:15 am, respectively, and ran for 50 minutes. Section C began at 12:20 pm and concluded at 1:10 pm. Because Section C occurred early in the afternoon, possibly just before or after lunch, the students may have been hungry or drowsy and thereby less motivated to engage and participate in class discussion.

The aforementioned critical thinking scores in the Section C class (approximately 10 percentile points higher than in the other sections) predicts that these students will likely have higher achievement levels, reflected by exam scores in the course (Williams, Oliver et al., 2003; Williams, Oliver et al., 2004). Once again, this typical relationship between critical thinking scores and exam performance was not reflected by Section C's performance on the unit exams. Tables 12, 13 and 14 show the exam scores for all sections of the course across all of the units within groups. Sections A and B consistently outperformed Section C on the unit exams as evidenced by group means. Despite this unexpected pattern, the positive correlation between exam scores and critical thinking skills was verified by significant positive correlations between exam scores and critical thinking scores across all sections of the course (see Table 11). Overall, it appears that students in Section C of the course were not as motivated to obtain course credit as the other sections of the course as evidenced by generally lower unit exam scores and lower levels of participation.

Strengths and Limitations of the Current Study

The greatest strength of the current study is its creative use of cooperative-learning group principles. This research is the first to directly measure of the effect of cooperative-learning groups on class participation. More specifically, no previous research study has used levels of class participation to form cooperative groups and subsequently reward groups for increased participation of all group members.

One limitation in this study was the effectiveness of the credit contingencies across only two of the three sections. Section C was consistently unresponsive to both individual and group credit contingencies. While Sections A and B demonstrated that the combined contingencies of individual credit and group credit produced more class participation than no credit or individual credit only, Section C showed limited effects from the combined group and individual contingencies. As previously examined, one of the most likely reasons for Section C's lack of increased participation was the credit-producing levels of class participation during baseline.

Another limitation of the study is the small amount of time in which students had the opportunity to participate. The class periods lasted 50 minutes and, during this time, 55 students were asked to participate two times during class in order to receive full credit. A few students expressed concern about this arrangement at the beginning of the semester. However, no students in any section of the course complained to the instructor that they did not have time to comment or were not called on by the instructor to volunteer a comment. Furthermore, class teachers were instructed to call on students who had not had an opportunity to participate in class during that class period, while waiting to call on more verbal class members until other students had a chance to participate. However, while this method of calling on students allowed as many students as possible to participate in class, it may have artificially capped participation of the

most vocal students. In particular, there may have been more dominant students identified in the class if the teachers had called on students randomly, without regard to their previous amount of participation. Counting raised hands to participate may be a more accurate way of assessing student engagement.

Variability across the three instructors may have introduced additional extraneous variables in the study. In an ideal study, the same instructor would teach all three researched sections of a course. In order to lessen the effect of instructional variation, several procedures were followed. First, all teachers were given a list of questions to pose in class discussion. The list of questions consisted of comprehension questions to ask students to ensure that they had a strong understanding of the course material. Second, all teachers were provided with instruction on teaching by the same faculty mentor. This faculty mentor ensured that all of the instructors understood the comprehension questions and could adequately provide feedback to all of the students on the completeness and correctness of their answers.

The varying levels of difficulty across different units were an additional limitation to this study. The first unit of the course discussed physical development of children and is a content area with which many of the students are familiar upon entering the class. Thus, participation may be more frequent in this section because students feel more comfortable volunteering comments when they are more acquainted with the content area. Conversely, the second unit of the course tends to be a more difficult unit of the class because it deals with cognitive issues previously unfamiliar to students. Finally, the last unit of the course (Unit 5) deals with more controversial issues (such as political values). Course material on controversial topics may impede class discussion in some cases. Many undergraduates may be uncomfortable with

confrontation and choose not to participate in class in order to avoid conflict with other class members or for fear that their views may not be accepted by other students or by the instructor.

Perhaps the greatest limitation in this study was the failure of the group-credit contingencies to facilitate participation in many of the small groups. Due to the requirement of all students within a group to participate for the group to receive bonus credit, I expected that the high-participating students would provide assistance to the low- or non-participants by helping them formulate a comment or question about the course material before or during class. Although there is a small amount of anecdotal evidence that group members encouraged each other to participate, this pattern was limited only to the groups within Section A of the course.

Conclusions

The purpose of this research study was to use cooperative learning principles to promote classroom participation in a large undergraduate course. Although results across the sections were mixed, the majority of class sections demonstrated that providing group-plus-individual credit for class participation resulted in higher levels of participation than during non-credit units. Furthermore, combining individual and group credit produced significantly higher rates of class participation than providing only individual credit. However, as demonstrated in one section of the course, providing credit for class participation did not appear to be beneficial when many of the students were already participating at credit-levels or higher. Thus, the most definitive conclusion from the findings is that providing participation credit, whether individual or individual-plus-group credit, will help initially-reticent students participate more in class discussion. However, combining group and individual participation credit will have more positive effects on participation than providing only individual credit for participation. Finally,

the effects of participation credit on the participation of initially-low participants is likely to be more pronounced earlier than later in a course.

List of References

Aspiranti, K. B. (2010). *Effects of random and delayed participation credit on participation levels in large college classes.* Unpublished doctoral dissertation. The University of Tennessee, Knoxville.

Aspiranti, K. B., McCleary, D. F., Foster, L. N., Galyon, C. E., Blondin, C. A., Yaw, J. S., & Williams, R.L. (in press). *Participation percentages under random and delayed credit contingencies. Journal of Excellence in College Teaching.*

Bean, J. C., & Peterson, D. (1998). Grading classroom participation. *New Directions in Teaching and Learning, 74*, 33-40.

Boniecki, K. A., & Moore, S. (2003). Breaking the silence: Using a token economy to reinforce classroom participation. *Teaching of Psychology, 30*, 224-227.

Bowen, C. W. (2000). A quantitative literature review of cooperative learning effects on high school and college chemistry achievement. *Journal of Chemical Education, 77*, 116 – 119.

Burchfield, C. M., & Sappington, J. (1999). Participation in classroom discussion. *Teaching of Psychology, 26*, 290-291.

Connor-Greene, P. A. (2005). Fostering meaningful classroom discussion: Student-generated questions, quotations, and talking points. *Teaching of Psychology, 32,* 173-175.

Fassinger, P. A. (1995a). Professors' and students' perceptions of why students participate in class. *Teaching Sociology*, *24,* 25-33.

Fassinger, P. A. (1995b). Understanding classroom interaction: Students' and professors' contribution to students' silence. *The Journal of Higher Education, 66,* 82-96.

Fassinger, P. A. (2000). How classes influence students' participation in college classrooms. *Journal of Classroom Interaction, 35,* 38-47.

Foster, L. N., Krohn, K. R., McCleary, D. F., Miller, K. B., Nalls, M. L., Quillivan, C. C., Taylor, C. M., & Williams, R. L. (2009). Increasing low-verbal students' participation in class discussion. *Journal of Behavioral Education, 18,* 173-188.

Fritschner, L. M. (2000). Inside the undergraduate college classroom: Faculty and students differ on the meaning of student participation. *The Journal of Higher Education, 71,* 342-362.

Galyon, C. E., Blondin, C. A., Forbes, B. E., & Williams, R. L. (in press). Does homework matter? A comparison of homework with established predictors of exam performance in large college classes. *Journal of Excellence in College Teaching*.

Hassan, K. E., & Madhum, G. (2007). Validating the Watson Glaser Critical Thinking Appraisal. *Higher Education, 54*, 361-383.

Hodge, G. K., & Nelson, N. H. (1991). Demonstrating differential reinforcement by shaping classroom participation. *Teaching of Psychology, 18,* 239-241.

Howard, J. R., Short, L. B., & Clark, S. M. (1996). Students' participation in the mixed-age college classroom. *Teaching Sociology, 23,* 8-24.

Jalilifar, A. (2010). The effect of cooperative learning techniques on college students' reading comprehension. *System: An International Journal of Educational Technology and Applied Linguistics, 38*, 96-108.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational Psychology Review*, *19*, 15-29.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998). Cooperative learning returns to college: What evidence is there that it works? *Change*, *30*(*4*), 26-35.

Krohn, K. R., Aspiranti, K. B., Foster, L. N., McCleary, D. F., Taylor, C. M., Nalls, M. L., Quillivan, C. C., & Williams, R. L. (2010). Effects of self-recording and contingent credit on balancing participation across students. *Journal of Behavioral Education, 19*, 134-155.

Krohn, K. R., Foster, L. N., McCleary, D. F., Aspiranti, K. B., Nalls, M. L., Quillivan, C. C., Taylor, C. M., & Williams, R. L. (2011). Reliability of students' self-recording participation in class discussion. *Teaching of Psychology*, *38*, 43-45.

Law, Y. (2008). Effects of cooperative learning on second graders' learning from text. *Educational Psychology, 28,* 567-582.

Ma, H. H. (2006). An alternative method for quantitative synthesis of single-subject researches: Percentage of data points exceeding the median. *Behavior Modification, 30*, 598-617.

McCleary, D. F., Aspiranti, K. B., Foster, L. N., Blondin, C. A., Galyon, C. E., Yaw, J. S., Forbes, B. N., & Williams, R. L. (2011). Balancing participation across students in a large college courses via randomized participation credit. *Journal of General Education, 64*, 194-214.

Shimazoe, J., & Aldrich, H. (2010). Group work can be gratifying: Understanding and overcoming resistance to cooperative learning. *College Teaching, 58*, 52-57.

Slavin, R. E. (1983). When does cooperative learning increase student achievement? *Psychological Bulletin, 94*, 429-445.

Slavin, R. E. (1985). An introduction to cooperative learning research. In R.E. Slavin, S. Sharan, S. Kagan, R. H. Lazarowitz, C. Webb, & R. Schmuck (Eds.) *Learning to cooperate, cooperating to learn* (pp. 5-15). New York, NY: Plenum Press.

Slavin, R. E. (1991). Synthesis of research on cooperative learning. *Educational Leadership*, *48*(5), 71-82.

Slavin, R. E. (2011). Instruction based on cooperative learning. In. R.E. Mayer (Ed.), *Handbook of research on learning and instruction* (pp. 344-360). New York, NY: Routledge.

Smith, D. G. (1977). College classroom interactions and critical thinking. *Journal of Educational Psychology, 69*(*2*), 180-190.

Terenzini, P. T., Cabrera, A. F., Colbeck, C. L., & Bjoklund, S. A. (2001). Collaborative learning vs. lecture/discussion: Students reported learning gains. *Journal of Engineering Education, 90*(1), 123-130.

Tsui, L. (1999). Courses and instruction affecting critical thinking. *Research in Higher Education, 40*, 185-200.

Turner, H. C., Bliss, S. L., Hautau, B., Carroll, E., Jaspers, K. E., & Williams, R. L. (2006). Brief daily writing activities and performance on major multiple-choice exams. *Journal of General Education, 55*, 221 – 246.

Watson, G. B., & Glaser, E. M. (1994). *Watson-Glaser critical thinking appraisal Form S manual*. San Antonio ,TX: Harcourt Brace.

Williams, R.L., & Clark, L. (2004). College students' ratings of student effort, student ability and teacher input as correlates of student performance on multiple-choice exams. *Educational Research, 46,* 229-239.

Williams, R. L., Oliver, R., Allin, J. L., Winn, B., & Booher, C. S. (2003). Psychological critical thinking as a course predictor and outcome variable. *Teaching of Psychology, 30,* 220-223.

Williams, R. L., Oliver, R., & Stockdale, S. (2004). Psychological versus generic critical thinking as predictors and outcome measures in a large undergraduate human development course. *The Journal of General Education, 53*, 37-58.

Williams, R. L., & Stockdale, S. L. (2003). High-performing students with low critical thinking skills. *The Journal of General Education, 52*, 200-226.

Wolery, M., Busick, M., Reichow, B., & Barton, E. E. (2010). Comparison of overlap methods for quantitatively synthesizing single-subject data. *The Journal of Special Education, 44*, 18-28.

Appendices

Appendix A: Tables

Table 1

Visual Depiction of Treatment Sequence

	Units				
	1	2	3	4	5
Section A	Baseline	IC	IC/GC	IC	IC/GC
Section B	Baseline	IC/GC	IC	IC/GC	IC
Section C	Baseline	IC/GC	IC	IC/GC	IC

Note. IC = Individual Credit, GC = Group Credit

Table 2

Percent Agreement between Student and Observer Records of Participation

	Units				
Reliability Pairs	1	2	3	4	5
Section A	b	IC	IC/GC	IC	IC/GC
Students and observer	67	84	75	80	71
Section B	b	IC/GC	IC	IC/GC	IC
Students and observer	92	81	70	80	80
Section C	b	IC/GC	IC	IC/GC	IC
Students and observer	74	78	81	67	72
All Sections					
Students and observer	78	81	73	74	74

Note. [b] = baseline; [IC] = individual credit only; [IC/GC] = individual and group credit

Table 3

Means and Standard Deviations for Student and Observer Records of Class Participation on the Inter-rater Check Day in Each Unit

	Units									
	Unit 1		Unit 2		Unit 3		Unit 4		Unit 5	
Section	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Section A	[b]		IC		IC/GC		IC		IC/GC	
Students	1.24	(1.24)	1.16	(1.13)	1.89	(0.86)	1.29	(0.98)	1.25	(0.81)
Observer 1	1.31	(1.22)	1.42	(1.27)	1.33	(0.93)	1.11	(1.07)	1.29	(1.01)
Section B	[b]		IC/GC		IC		IC/GC		IC	
Students	1.35	(1.27)	1.33	(0.903)	1.24	(1.10)	1.55	(0.97)	1.12	(1.01)
Observer 1	1.39	(1.48)	1.53	(1.12)	1.25	(1.09)	1.24	(1.20)	1.38	(1.24)
Section C	[b]		IC/GC		IC		IC/GC		IC	
Students	1.57	(0.91)	1.48	(1.18)	1.42	(0.99)	1.62	(1.05)	1.17	(0.93)
Observer 1	1.35	(0.93)	1.70	(1.93)	1.54	(1.17)	1.52	(1.19)	1.50	(1.32)
All Sections										
Students	1.41	(1.15)	1.57	(1.07)	1.47	(0.99)	1.30	(1.00)	1.39	(0.96)
Observer 1	1.35	(1.23)	1.54	(1.46)	1.37	(1.06)	1.28	(1.16)	1.39	(1.19)

Note. [b] = baseline; [IC] = individual credit only; [IC/GC] = individual and group credit

Table 4

Number of Questions (Factual, Comprehension, Total) Posed by Instructors

	Units				
	1	2	3	4	5
Section A	[b]	IC	IC/GC	IC	IC/GC
Factual	9	34	13	13	11
Comprehension	17	22	24	19	19
Total	26	56	37	32	30
Section B	[b]	IC/GC	IC	IC/GC	IC
Factual	15	16	6	11	14
Comprehension	15	17	14	21	14
Total	30	33	20	32	28
Section C	[b]	IC/GC	IC	IC/GC	IC
Factual	1	19	12	17	12
Comprehension	11	20	25	19	22
Total	12	39	37	36	34

Note. [b] = baseline; [IC] = individual credit only; [IC/GC] = individual and group credit

Table 5

Percent Agreement between Observers on Factual, Comprehension and Total Questions

	Units				
Reliability Pairs	1	2	3	4	5
Section A	[b]	IC	IC/GC	IC	IC/GC
Factual Questions	78	82	81	68	92
Comprehension Questions	89	76	83	47	100
Total Questions	100	98	97	87	97
Section B	[b]	IC/GC	IC	IC/GC	IC
Factual Questions	100	88	75	79	79
Comprehension Questions	73	41	93	67	100
Total Questions	87	64	95	79	89
Section C	[b]	IC/GC	IC	IC/GC	IC
Factual Questions	100	79	50	71	83
Comprehension Questions	100	77	78	61	81
Total Questions	100	95	97	84	92

Note. [b] = baseline; [IC] = individual credit only; [IC/GC] = individual and group credit

Table 6

Percent Agreement between Observer Records of Positive and Negative Teacher Feedback

	Units				
Reliability Pairs	1	2	3	4	5
Section A	[b]	IC	IC/GC	IC	IC/GC
Positive Feedback	100	99	97	100	95
Negative Feedback	0	100	0	100	33
Section B	[b]	IC/GC	IC	IC/GC	IC
Positive Feedback	74	94	90	98	96
Negative Feedback	100	80	100	50	100
Section C	[b]	IC/GC	IC	IC/GC	IC
Positive Feedback	98	95	100	97	96
Negative Feedback	100	0	0	0	100

Note. [b] = baseline; [IC] = individual credit only; [IC/GC] = individual and group credit

Table 7

Mean Daily Participation Levels Based on Student Daily Records

	Day 1	Day 2	Day 3	Day 4	Overall
Unit 1					
Section A	1.31	1.42	1.04	1.24	1.25
Section B	1.27	1.39	1.44	1.35	1.36
Section C	1.72	1.44	1.27	1.57	1.50
Unit 2					
Section A	1.57	1.45	1.54	1.16	1.43
Section B	1.62	1.70	1.84	1.33	1.63
Section C	1.46	1.53	1.80	1.48	1.57
Unit 3					
Section A	1.60	1.45	1.56	1.89	1.63
Section B	1.70	1.28	1.73	1.24	1.49
Section C	1.66	1.69	1.35	1.42	1.53
Unit 4					
Section A	1.30	1.23	1.40	1.29	1.30
Section B	1.61	1.24	1.87	1.55	1.57
Section C	1.46	1.48	1.45	1.62	1.50
Unit 5					
Section A	1.27	1.35	1.16	1.25	1.26
Section B	1.31	1.39	1.47	1.12	1.32
Section C	1.35	1.44	1.46	1.17	1.36

Table 8

ANOVA Table for Sections, Contingencies, and Application Times for All Students

Source of Variation	*df*	SS	MS	*F* ratio	*p*
Between Subjects					
Section	2	26.930	13.465	.376	.688
Within Subjects					
Application	1	85.452	85.452	31.448	.001
Contingency	1	20.779	20.779	6.117	.014
Application by Contingency	1	.286	.286	.125	.724
Application by Section	2	5.342	2.671	.983	.377
Contingency by Section	2	8.648	4.324	1.273	.283
Application by Contingency by Section	2	7.436	3.718	1.623	.201

Note. SS = sums of squares; *df* = degrees of freedom; MS = mean squares

Table 9

ANOVA Table for Sections, Contingencies, and Application Times for Initially-Low Participants

Source of Variation	*df*	SS	MS	*F* ratio	*p*
Between Subjects					
Section	2	13.649	13.649	.745	.481
Within Subjects					
Application	1	11.255	11.255	4.833	.034
Contingency	1	30.251	30.251	8.319	.006
Application by Contingency	1	.013	.013	.008	.930
Application by Section	2	7.493	3.746	1.609	.213
Contingency by Section	2	21.639	10.819	2.975	.062
Application by Contingency by Section	2	17.883	8.942	5.389	.008

Note. SS = sums of squares; *df* = degrees of freedom; MS = mean squares

Table 10

Correlations Between Critical Thinking Raw Scores and Participation in Each Unit

	Units				
	1	2	3	4	5
Section A	.21	.14	.39*	.31*	.14
Section B	.43**	.24	.21	.18	.21
Section C	-.19	.03	-.06	-.07	-.04

Note: *indicates correlations that are significant at the .05 level; ** indicates correlations that are significant at the .01 level

Table 11

Correlations Between Critical Thinking Raw Scores and Unit Exam Scores

	Units				
	1	2	3	4	5
Section A	.23	.36**	.33*	.41**	.32*
Section B	.46**	.48**	.26	.40**	.42**
Section C	.20	.25	.26	.19	.32*

Note: *indicates correlations that are significant at the .05 level; ** indicates correlations that are significant at the .01 level

Table 12

Means and Standard Deviations of Section A Group Achievement Levels on Unit Exams

	Units				
	1	2	3	4	5
	[b]	IC	IC/GC	IC	IC/GC
Group 1	41.60 (4.51)	37.20 (6.14)	40.80 (4.09)[2]	39.00 (4.90)	37.80 (3.27)[0]
Group 2	37.40 (3.21)	36.00 (3.32)	38.60 (3.36)[1]	39.60 (4.67)	39.00 (4.36)[0]
Group 3	36.40 (2.70)	36.20 (5.31)	38.60 (6.23)[0]	36.00 (4.32)	39.50 (10.54)[2]
Group 4	39.60 (3.36)	42.00 (3.61)	41.80 (3.03)[0]	39.80 (4.44)	40.20 (3.90)[0]
Group 5	37.40 (3.85)	33.80 (9.15)	34.60 (4.83)[2]	35.00 (6.44)	38.00 (7.42)[2]
Group 6[2]	40.20 (3.96)	42.40 (3.36)	41.00 (3.39)[1]	41.00 (2.00)	43.40 (3.21)[1]
Group 7[4]	39.00 (4.06)	37.20 (8.53)	38.80 (6.06)[2]	39.80 (4.21)	36.80 (12.05)[2]
Group 8[4]	41.20 (4.03)	42.00 (5.61)	40.60 (8.35)[2]	41.20 (9.42)	42.60 (5.03)[2]
Group 9[1]	42.40 (5.94)	40.20 (3.35)	41.60 (4.83)[1]	42.40 (5.64)	42.80 (4.61)[0]
Group 10[3]	42.00 (5.80)	39.00 (6.69)	41.67 (4.27)[2]	43.50 (5.01)	41.17 (5.98)[1]
Group 11[4]	43.80 (4.38)	36.20 (6.02)	42.00 (4.06)[2]	42.60 (6.77)	40.40 (5.68)[2]

Note. [b] = baseline; [IC] = individual credit only; [IC/GC] = individual and group credit; [0] = group received no bonus credit for group participation during the specified unit; [1] = group received partial bonus credit for group participation during the specified unit; [2] = group received full bonus credit for group participation during the specified unit.

Table 13

Means and Standard Deviations of Section B Group Achievement Levels on Unit Exams

	Units				
	1	2	3	4	5
	[b]	IC/GC	IC	IC/GC	IC
Group 12	42.80 (3.27)	42.40 (5.93) [0]	43.40 (2.97)	43.40 (3.05) [1]	42.20 (6.10)
Group 13 [1]	40.80 (6.14)	43.20 (4.87) [0]	45.00 (2.55)	44.20 (2.59) [1]	42.20 (3.70)
Group 14 [1]	44.80 (3.27)	43.80 (4.87) [1]	44.00 (2.45)	45.40 (3.85) [0]	44.00 (3.94)
Group 15	39.00 (6.48)	40.40 (5.98) [0]	40.80 (4.92)	42.60 (5.41) [0]	42.00 (4.18)
Group 16	39.20 (5.50)	37.00 (9.25) [0]	41.00 (2.24)	39.40 (6.62) [0]	39.20 (7.23)
Group 17	41.20 (2.59)	38.60 (6.15) [2]	41.20 (3.56)	41.80 (4.66) [2]	40.40 (3.85)
Group 18	41.00 (5.15)	37.80 (6.46) [0]	41.20 (3.03)	41.20 (4.32) [0]	40.60 (1.95)
Group 19	42.60 (4.28)	45.20 (4.97) [0]	40.60 (3.36)	40.40 (8.36) [0]	42.60 (6.19)
Group 20 [2]	39.20 (3.96)	35.40 (7.30) [0]	36.40 (7.44)	36.00 (4.53) [2]	36.00 (6.48)
Group 21	43.20 (5.26)	44.20 (3.77) [0]	42.80 (2.17)	45.60 (3.78) [0]	42.40 (5.60)
Group 22 [2]	38.20 (3.49)	36.40 (5.32) [1]	37.40 (7.23)	37.40 (6.91) [1]	43.20 (4.76)

Note. [b] = baseline; [IC] = individual credit only; [IC/GC] = individual and group credit; [0] = group received no bonus credit for group participation during the specified unit; [1] = group received partial bonus credit for group participation during the specified unit; [2] = group received full bonus credit for group participation during the specified unit.

Table 14

Means and Standard Deviations of Section C Group Achievement Levels on Unit Exams

	Units				
	1	2	3	4	5
	[b]	IC/GC	IC	IC/GC	IC
Group 23	36.20 (5.54)	35.60 (4.98) [0]	40.60 (2.61)	41.00 (2.45) [0]	43.00 (3.00)
Group 24	39.50 (5.68)	32.00 (10.14) [0]	36.67 (4.76)	39.17 (8.80) [0]	42.33 (3.50)
Group 25	36.20 (3.63)	37.00 (2.83) [0]	38.50 (3.87)	40.75 (4.57) [0]	43.00 (4.08)
Group 26	38.20 (5.59)	36.50 (2.08) [0]	38.40 (3.21)	39.50 (6.46) [0]	41.20 (4.55)
Group 27	37.20 (2.17)	30.00 (7.35) [0]	40.00 (4.36)	37.40 (8.11) [0]	39.00 (1.87)
Group 28	33.80 (6.38)	39.80 (4.03) [0]	42.00 (1.41)	44.60 (3.05) [2]	41.20 (7.23)
Group 29	36.60 (4.34)	35.40 (2.41) [0]	39.80 (6.46)	37.00 (9.27) [0]	39.00 (5.83)
Group 30	38.40 (4.83)	41.60 (7.13) [0]	42.20 (5.76)	41.40 (6.43) [0]	44.60 (4.22)
Group 31	36.00 (2.00)	31.80 (8.96) [0]	39.40 (4.34)	41.20 (6.72) [2]	39.80 (5.76)
Group 32	40.00 (6.38)	34.40 (6.73) [0]	38.40 (6.03)	38.20 (4.21) [0]	43.00 (3.87)
Group 33	39.20 (7.40)	32.00 (11.02) [0]	39.20 (5.80)	37.20 (7.89) [0]	42.40 (4.83)

Note. [b] = baseline; [IC] = individual credit only; [IC/GC] = individual and group credit; [0] = group received no bonus credit for group participation during the specified unit; [1] = group received partial bonus credit for group participation during the specified unit; [2] = group received full bonus credit for group participation during the specified unit.

Appendix B: Figures

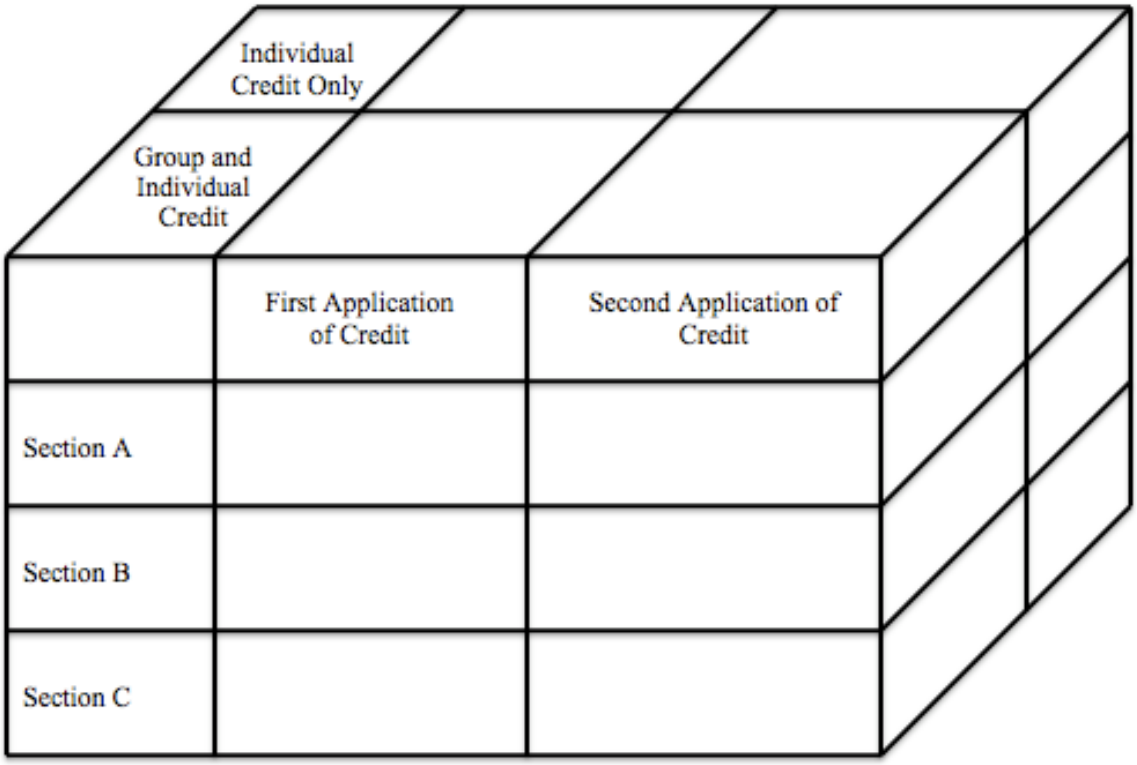

Figure 1. Diagram of mixed three-way ANOVA analysis.

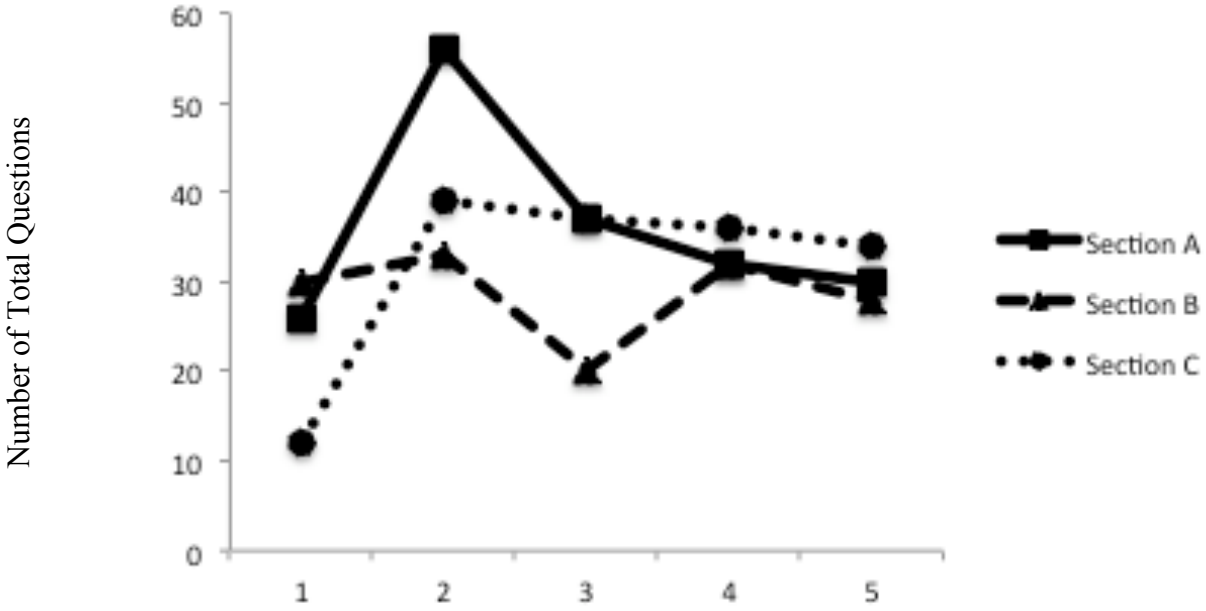

Figure 2. Number of total questions posed per unit by the instructor on inter-rater days.

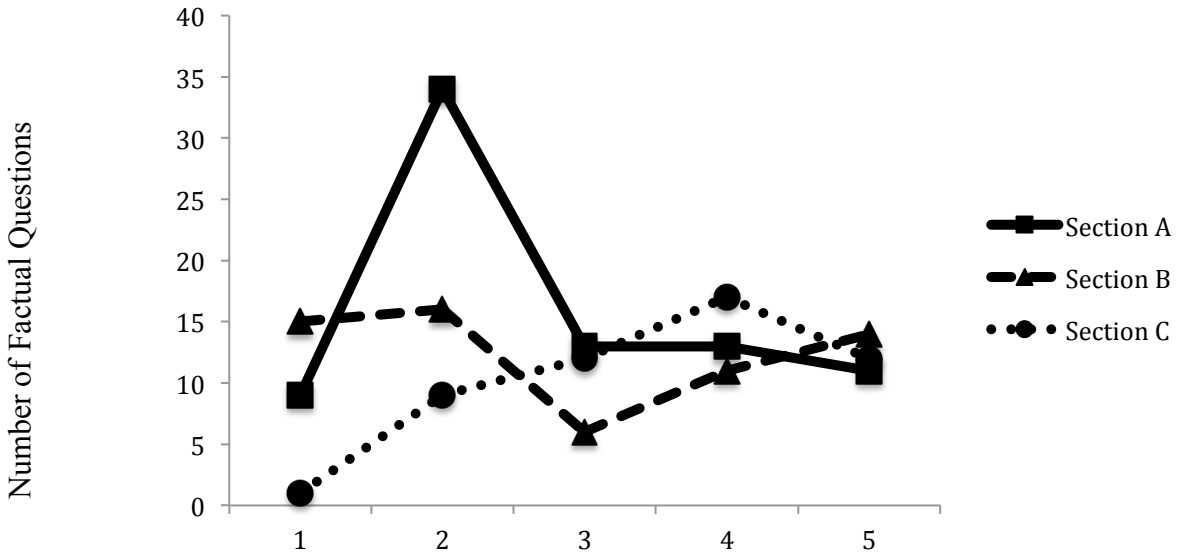

Figure 3. Number of factual questions posed by the instructor in each unit on inter-rater day.

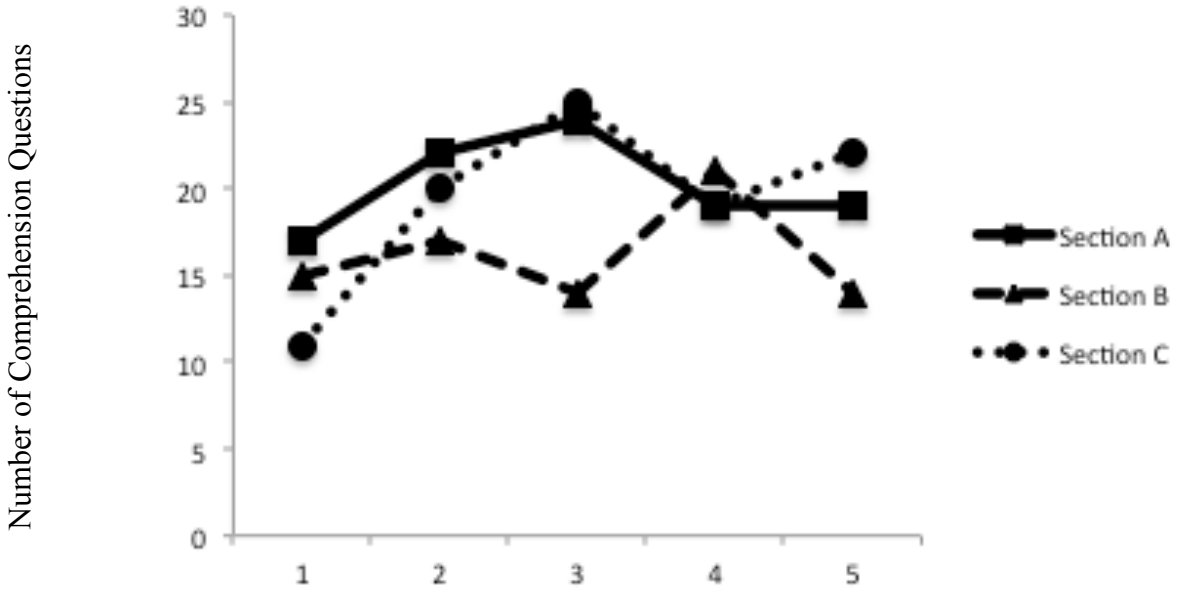

Figure 4. Number of comprehension questions posed by the instructor in each unit on inter-rater day.

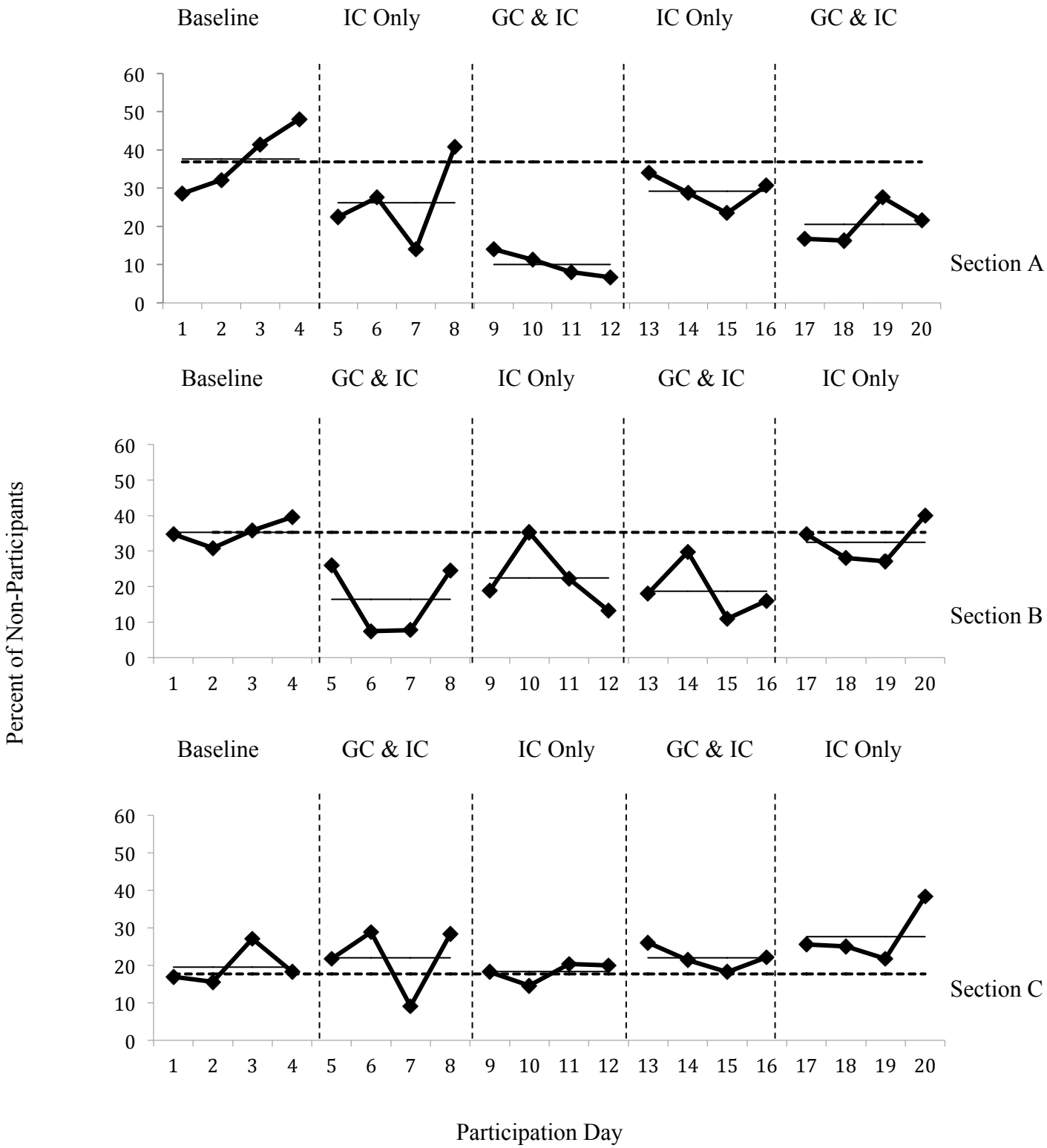

Figure 5. Percent of non-participants (0 comments) each day compared to a dotted line denoting the median percent of non-participants during baseline.

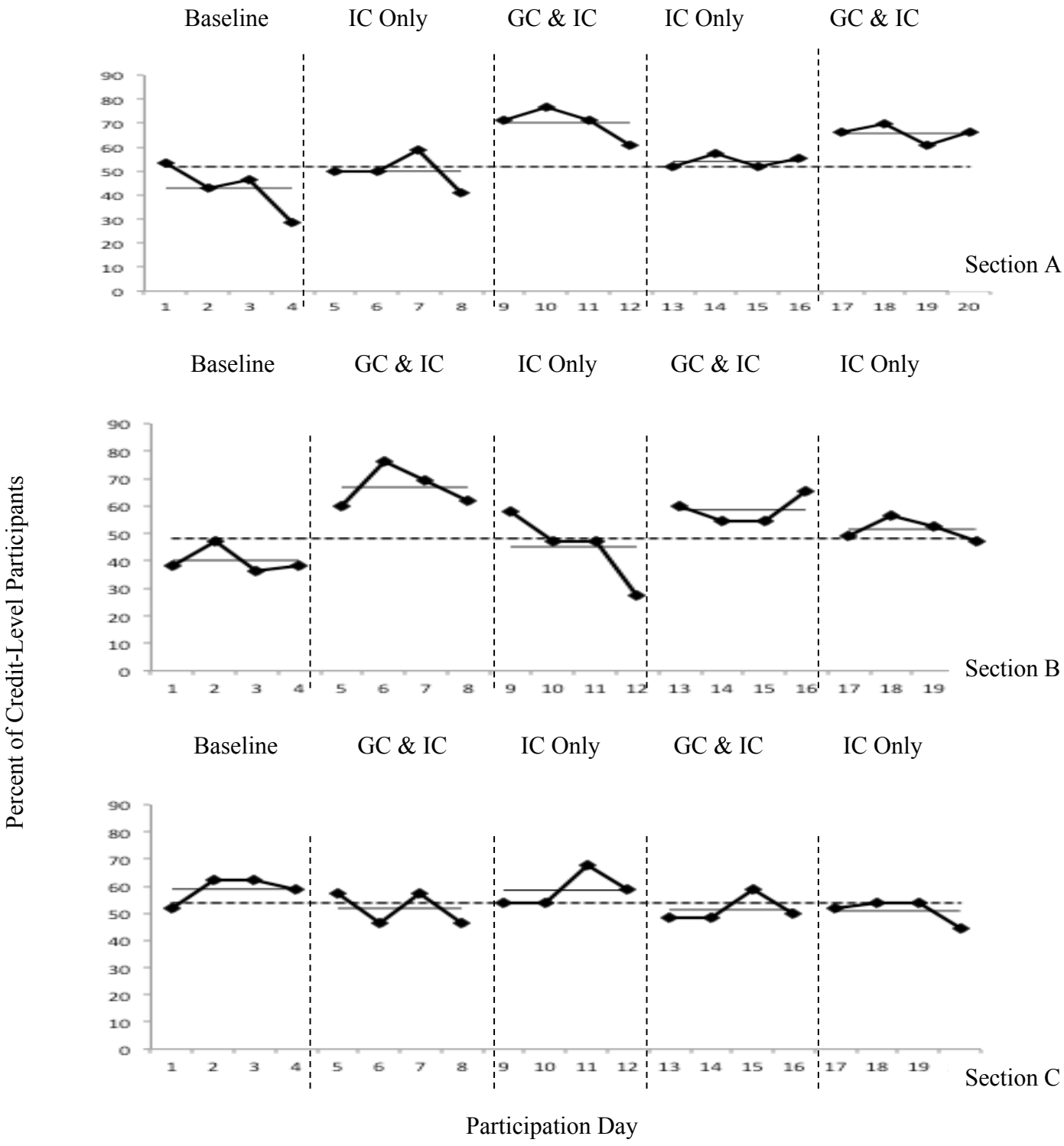

Figure 6. Percent of Credit-Level participants (1-2 comments) each day compared to a dotted line denoting the median percent of credit-level participants during the individual credit only phases.

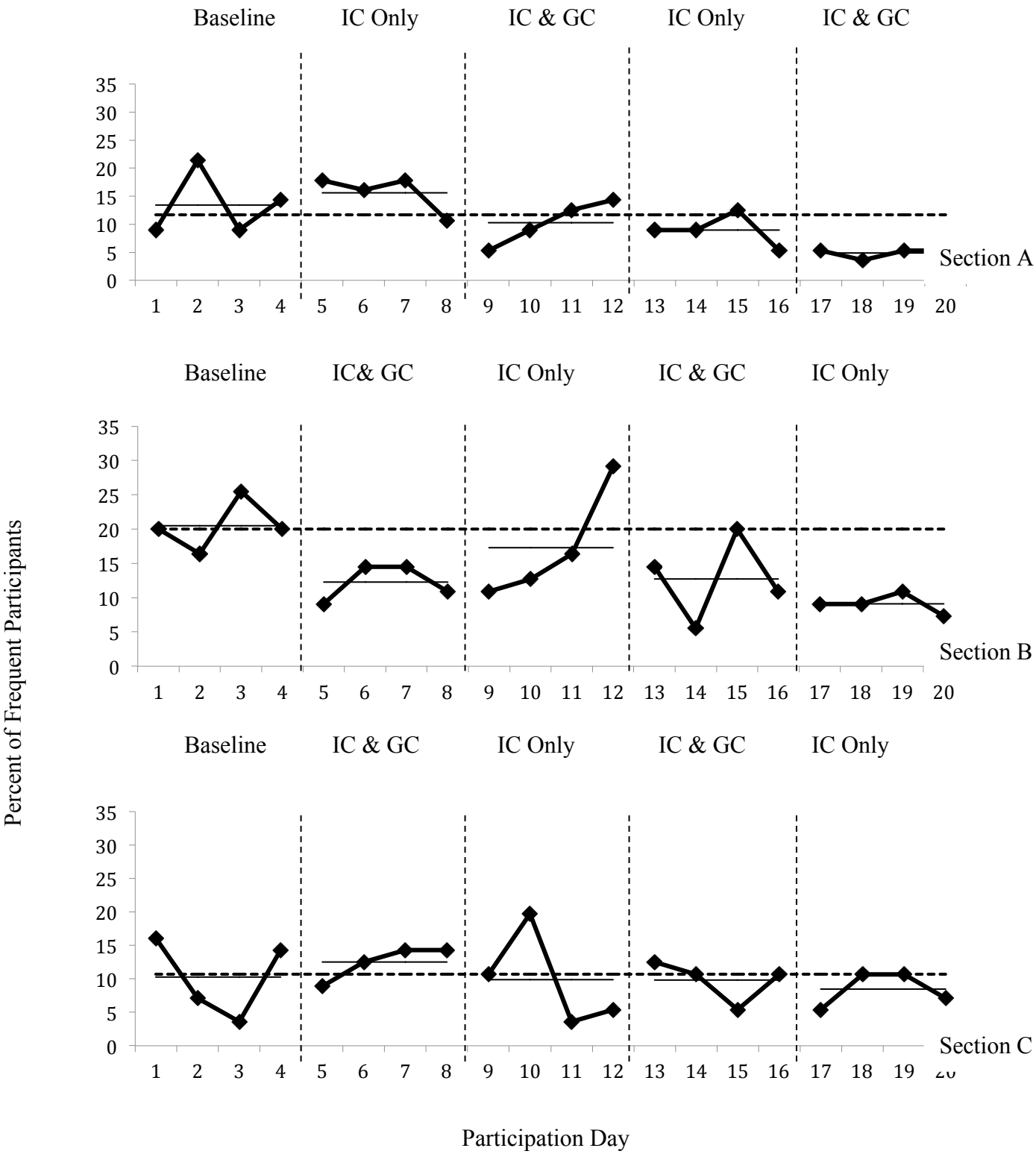

Figure 7. Percent of frequent participants (3-4 comments) each day compared to a dotted line denoting the median percent of frequent participants during the baseline phase.

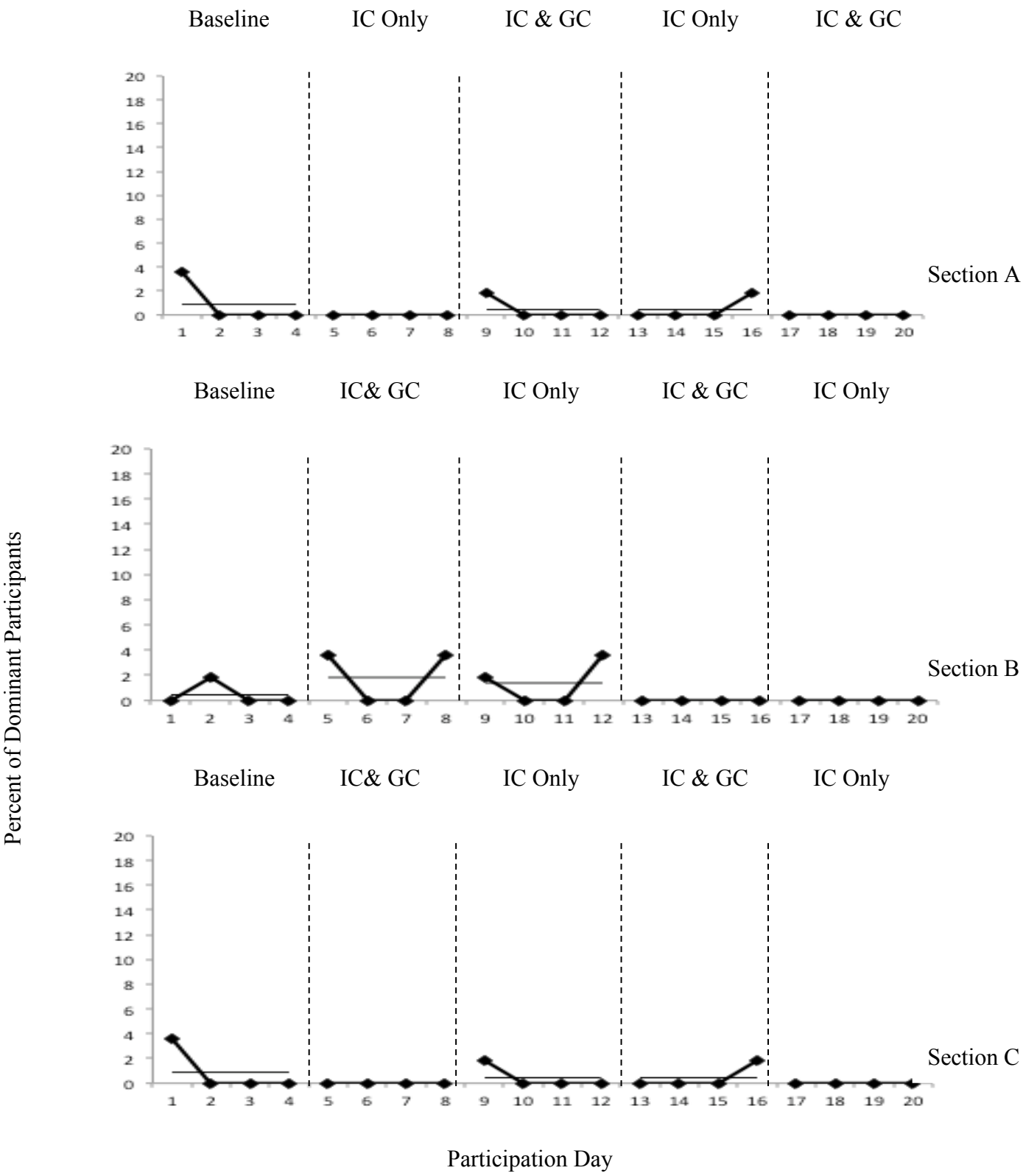

Figure 8. Percent of Dominant Participants (5+ comments) each day.

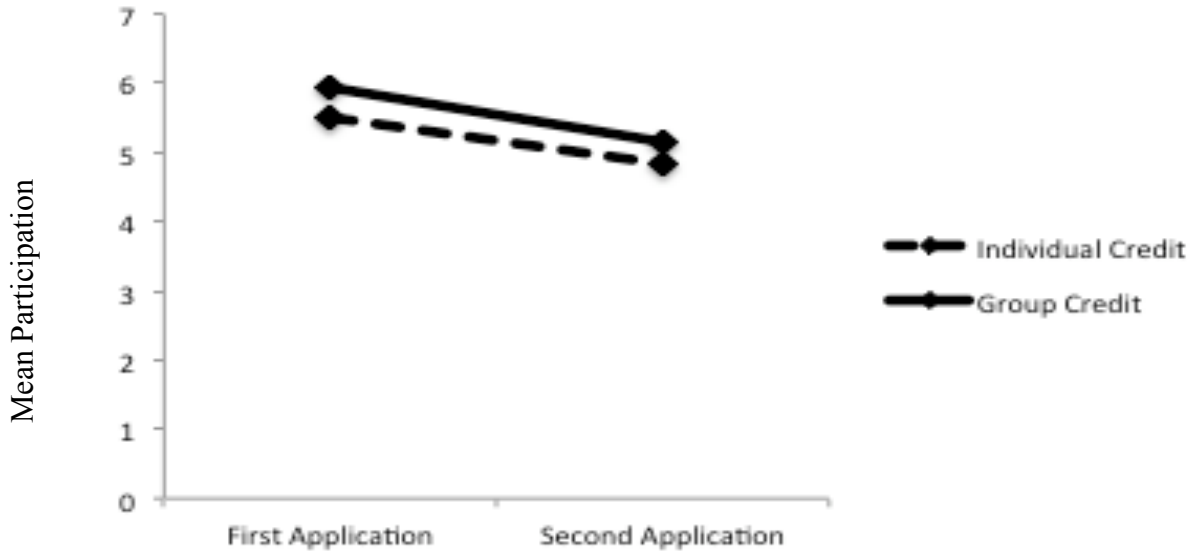

Figure 9. Mean participation during the first and second application of two types of credit contingencies.

Note: The interaction between credit and time of application was not significant. There was a significant main effect for credit contingency and a significant main effect for time of application.

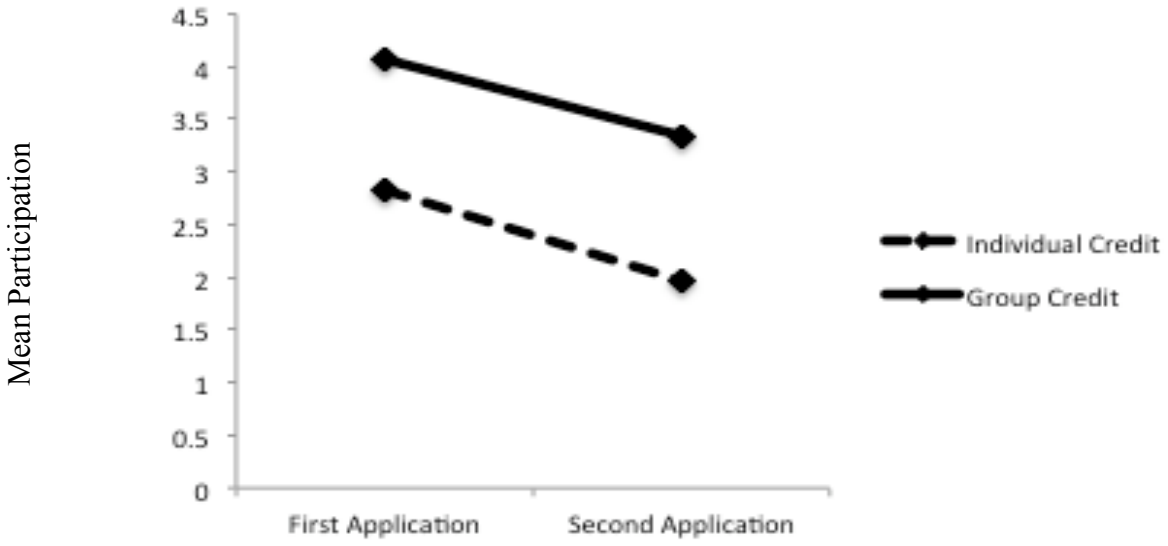

Figure 10. Mean participation during the first and second application of two types of credit contingencies for initially-low participants in Sections A and B

Note: A significant three-way interaction and significant main effects were obtained for both credit contingencies and time of application.

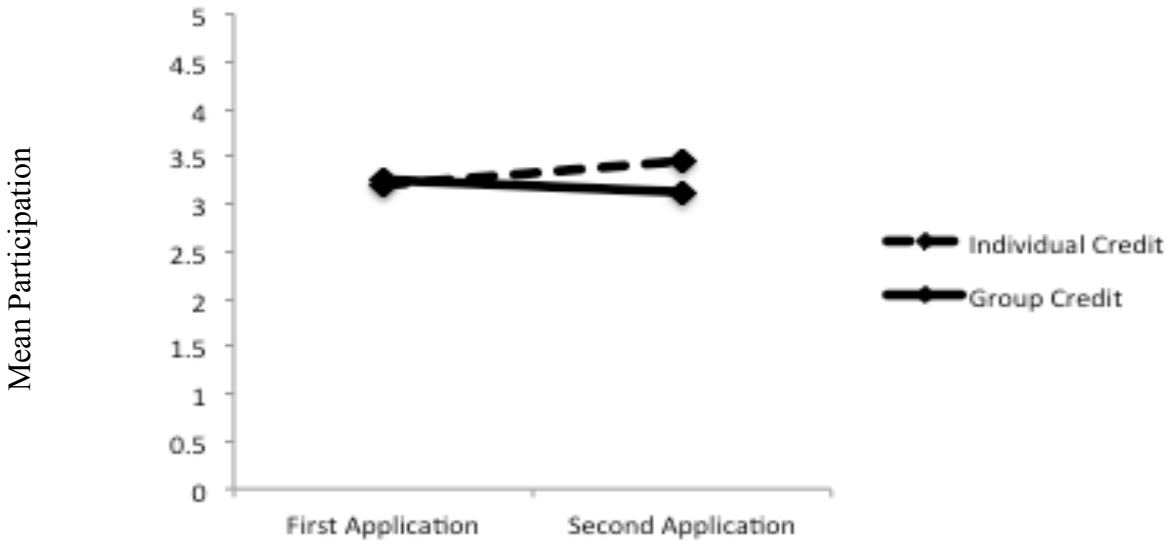

Figure 11. Mean participation during the first and second application of two types of credit contingencies for initially-low participants in Section C.

Note: Although a significant three-way interaction was obtained for section by contingency by time of application for initially-low participants, no significant simple effects were obtained in Section C for initially-low participants.

Vita

Cora Taylor was born and raised in St. Mary's County, Maryland. She graduated from Great Mills High School in Maryland in 2004 and moved to southwest Virginia to attend Radford University. She graduated from Radford University in 2007 with a Bachelors of Arts degree in Psychology. She continued her education at the University of Tennessee in Knoxville, entering the field of School Psychology. She graduated with a Masters of Science in Applied Educational Psychology in 2010 and a Doctor of Philosophy in School Psychology in 2012.