To the Graduate Council:

I am submitting herewith a dissertation written by Daniel Fox McCleary entitled "Balancing Student Participation in Large College Courses via Randomized Credit for Participation." 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:

David F. Cihak

R. Steve McCallum

Christopher H. Skinner

Accepted for the Council:

Carolyn R. Hodges
Vice Provost and Dean of the Graduate School

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

Balancing Student Participation in Large College Courses via Randomized Credit for Participation

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

Daniel Fox McCleary
December 2011

Copyright © 2011 by Daniel Fox McCleary

All rights reserved.

Acknowledgments

First, I would like to thank Dr. Robert Williams for his mentorship and guidance over the years. His work ethic, precision of writing, teaching ability, and dedication to physical health are just some of his many worthy qualities I wish to emulate in my own life. Second, I would like to thank my dissertation committee whose assistance throughout this process was essential. Dr. Chris Skinner provided me with many insights into the inner workings of the field of school psychology and the publication process. Dr. R. Steve McCallum and Dr. David Cihak whose incisive critiques have improved my understanding of statistics and writing in general. Third, I wish to thank Dr. Brian Wilhoit and Dr. Sherry Bain who have also been instrumental in training me as a school psychologist. Fourth, the data collection of this dissertation would not have been possible without the assistance of Lisa Foster, Carolyn Blondin, Charles Galyon, Jared Yaw, Bethany Forbes, Elizabeth Hall, and Jacob Goings. Fifth, all my professors and swim coaches who have helped shape me into the person I am today are all due a huge thank you. Last, a very special thank you is owed to Lisa Foster whose statistical guidance and writing assistance was indispensable.

Abstract

The current study was an extension of research reported by Krohn (2010), which showed that daily credit for self-reported participation in designated credit units tended to balance participation across students (i.e., fewer non-participants, more credit-level participants, and fewer dominant participants). The purpose of the current study was to determine if similar results would be achieved by randomly selecting half of the discussion days in designated credit units for participation credit.

The study was done in 3 large sections of an undergraduate class (approximately 54 students per class). Students self-recorded their in-class comments each day on specially designed record cards. At the end of each pre-selected unit, instructors randomly selected discussion days and awarded credit based on the number of comments made on the days randomly selected. Three credit points were given for each student's first comment and two additional points for a second comment.

The findings of the current study differed in several ways from those of Krohn's (2010) comparison study. The differences mainly related to baseline percentages of different levels of participation. Compared to the current study, Krohn's study had a higher percentage of non-participants, fewer credit-level participants, fewer frequent participants, and more dominant participants. The disparities between the baseline levels of Krohn's study and the current study made treatment effects more difficult to achieve in the latter study. Nonetheless, there were fewer non-participants and more credit-level and frequent participants during credit units than in non-credit units.

Secondarily, a survey was given at the beginning of the course to analyze student beliefs regarding participation. Using the same survey, Krohn (2010) extracted three primary factors: 1) Personal Benefits of Participation, 2) Expectation for Discussion in College Classes, and 3) Personal History and Confidence Regarding Participation. The same three factors were also examined separately and in combination in the current study. Results showed the three-factor model to predict student participation levels better than the total survey. In addition, students were given the Watson-Glaser Critical Thinking Appraisal at the onset of the course. A logistic regression indicated that exam and critical thinking scores, in combination, significantly predicted student participation levels.

Table of Contents

Chapter I 1
Introduction and Literature Review 1
Student Variables Affecting Participation 2
Instructor Variables Affecting Student Participation 3
Grading Student Participation 5
Assessing Student Participation 8
Participation History and Beliefs 10
Critical Thinking and Participation 11
Framework for the Current Study 12
Research Questions 15
Chapter II 16
Method 16
Participants 16
Environmental Context for the Study 16
Recording Procedures 18
Crediting Procedures 21
Survey of Student Participation Perspectives 22
Critical Thinking and Participation 23
Participation Contingencies and Research Design 24
Chapter III 25
Results 25
Descriptive Demographic Data and Correlations 26
Reliability Assessments 26
Effect of Credit Treatment on Levels of Student Participation 30
Low- and Non-Participants 35
Medium- and High-Participants 36
Timely and Repetitious Student Comments 37
Interaction between Treatment and Instructor Behavior 38
Survey Results 40
Critical Thinking Results 46
Chapter IV 49
Discussion and Conclusion 49
Effect of Credit on Patterns of Student Participation 49
Effect of Credit on Low- Medium- and High Participants 52
Reliability and Student Participation 53
Non-effect of Differential Instructor Behavior 53
Effect of Participation Credit on Quality of Comments 55
Participation Survey as a Predictor of Participation 57
Critical Thinking and Exam Grades as Predictors of Participation 58
Limitations and Future Directions 60
References 64
Appendices 72

Appendix A: Tables 73
Appendix B: Figures 96
Appendix C: Student Discussion Form 108
Appendix D: Instructor Discussion Form 109
Appendix E: Students' Under- and Over-Reporting of Class Participation during Non-Credit vs. Credit Units 110
Appendix F: Participation Survey 117
Appendix G: Means for Low, Medium, and High-Responding Participants (based on Unit 1 participation levels) on each Item on the Participation Survey 129
Vita 135

List of Tables

Table 1. Credit Contingency per Unit 74

Table 2. Mean Daily Participation Level Based on Student Gender and Academic Classification 75

Table 3. Mean percent of Inter-observer Agreement between Student and observer Records of Class Participation on Inter-rater Check Days 76

Table 4. Means and Standard Deviations for Student and Observer records of Class Participation on the Inter-rater Check Day in Each Unit 77

Table 5. Mean Percent of Inter-Observer Agreement between Student and Observer Records of Timely Student Participation on Inter-rater Check Days 78

Table 6. Mean Percent of Inter-Observer Agreement between Student and Observer Records of Repetitious Student Commenting on Basseline (b), Credit (c), and Non-credit (nc) Units 79

Table 7. Percent Agreement between Observer Records of Frequency of Instructor Questions on the Inter-rater Check Days in Each Unit 80

Table 8. Mean Percent of Inter-Observer Agreement between Observers' Records of Instructor Feedback to Students on Inter-rater Check Days in Each Unit 81

Table 9. Differences in Proportions of Students at Different Participation Levels between Treatment Conditions Overall and between pairs of Adjacent Treatment Units in Each Section 82

Table 10. Number of Low Participants in Unit 1 Who Fell into Low, Medium, and High Categories in Subsequent Units 84

Table 11. Number of Medium Participants in Unit 1 Who Fell into Low, Medium, and High Categories in Subsequent Units 85

Table 12. Number of High Participants in Unit 1 Who Fell into Low, Medium, and High Categories in Subsequent Units 86

Table 13. Number of Timely and Repetitious Comments per Unit Recorded by Students and Observers 87

Table 14. Number of Questions (Factual, Comprehension, Total) Posed by Instructors by Unit within Each Section 88

Table 15. Means and Standard Deviations for Positive and Negative Instructor Feedback following Student Participation per Inter-rater Check Day in Each Unit 89

Table 16. Logistic Regression Results using the Total Participation Survey Items to Predict Placement in Low and High Participation Groups each Unit 90

Table 17. Percent of Cases Correctly Classified into Low and High Groups each Unit by Section with the Total Participation Survey Score for Each Student as the Predictor Variable 91

Table 18. Logistic Regression Results using the Participation Survey Factors as Predictors in each Unit 92

Table 19. Percent of Cases Correctly Classified into Low and High Groups for the Combined Three Survey Factors as a Model for Predicting Participation 93

Table 20. Logistic Regression Results using the History/Confidence Participation Survey Factor as the Predictor in each Unit 94

Table 21. Percent of Cases Correctly Classified into Low and High Groups for the History/Confidence Survey Factor as a Model for Predicting Participation 95

List of Figures

Figure 1. Record card for recording participation and daily credit activities. 97

*Figure 2.*Percent of non-participants (0 comment level) each day. 98

Figure 3. Percent of Credit-Level Participants per day (1-2 comments only) in each unit. 99

Figure 4. Percent of frequent participants (3-4 comment level) each day..... 100

Figure 5. Percent of dominant participants (5+ comment level) each day. 101

Figure 6. Percent of low-responding participants in Unit 1 who participated in subsequent units. 102

Figure 7. Mean number of total questions per unit posed by the instructor in each section on inter-rater days. 103

Figure 8. Mean number of comprehension questions on inter-rater check days posed by the instructor in each unit..... 104

Figure 9. Percent of comprehension questions each inter-rater day posed by the instructor in each section. 105

Figure 10. Mean quantity of positive feedback provided to individual students within each unit on inter-rater check days. 106

Figure 11. Mean quantity of negative feedback provided to individual students within each unit on inter-rater check days. 107

Chapter I

Introduction and Literature Review

Educational researchers have devoted considerable attention to developing effective and efficient procedures for promoting class discussion (Aspiranti, 2010; Auster & MacRone, 1994; Boniecki & Moore, 2003; Connor-Greene, 2005; Fassinger, 1995a; Ferguson, 1986; Foster et al., 2009; Garside, 1996; Hodge & Nelson, 1991; Howard & Henney, 1998; Krohn, 2010; Krohn et al., 2010; Mainkar, 2008; Montello, 1988; Simpson & Courtney, 2008; Weaver & Qi, 2005). Bean and Peterson (1998) and Jones (2008) have documented that the overwhelming majority of college syllabi include participation as a course requirement. However, participation seldom is systematically recorded, especially in large undergraduate classes. Typically, participation is included in the syllabus in an effort to increase the amount of students' overt engagement in discussion and to encourage a broad array of students to discuss the topic being addressed (Boniecki & Moore; Howard, James, & Taylor, 2002). It is important to note that researchers often use the term "participation" to refer to verbal comments contributed during class discussion rather than covert engagement (e.g., thinking about course content and class discussion/lecture, taking notes) during class (Asprianti; Bean & Peterson; Foster et al.; Jones; Krohn; Krohn et al.).

Jones (2008) reported findings supporting the use of rewards/incentives for participation in class discussions. Results showed increased levels of attendance, increased preparation for class, and increased application of course knowledge to other course concepts. Others have found participation requirements to bolster accountability in class preparation, increase the number of students participating, increase exposure to divergent views, increase critical thinking

skills, and increase important communication skills (Bean & Peterson, 1998; Garside, 1996; Jones).

Student Variables Affecting Participation

Although participation researchers have found a myriad of benefits from increased participation, some students rarely, if ever, participate in class discussion. Some of the most common variables examined in relation to participation are identified below. Researchers have found conflicting results with regards to gender (Auster & MacRone, 1994; Fassinger, 1995b; Gilson, 1994; Howard & Henney, 1998; Howard et al., 2002). With respect to age, researchers have shown that students aged 25 and older are much more likely to engage in class discussions than students of traditional college age (Howard & Henney; Howard et al.). Shyness/introversion, lack of information/knowledge, and incomplete ideas are also frequently cited reasons for not participating in class (Christensen, Curley, Marquez, & Menzel, 1995; Connor-Greene, 2005; Fassinger; Howard & Henney; Howard et al.; Mainkar, 2008; Renne, Kass, & Nay, 1973; Tsui, 1998; Williams, Oliver, Allin, Winn, & Booher, 2003a; Weaver & Qi, 2005).

Regarding issues of shyness, Aitken and Neer (1993) found communicative apprehension to be a deterrent for only 20% of collegiate students. Fassinger (1995b) proposed lack of confidence, interest, and preparation to be likely contributors to one's lack of participation. Howard and Henney (1998) found one's perception of responsibility in class discussion to be a major factor in determining one's degree of participation. Also, they discovered that most class discussions derived from student-initiated questions or comments, indicating students took responsibility for class discussion. However, students' view of responsibility for class discussion

likely derives from the topic of discussion, instructor's manner of teaching, and from the class atmosphere established from the onset of the course.

Instructor Variables Affecting Student Participation

Although the importance of instructor behaviors affecting class participation has been anecdotally established, little empirically-based data support that claim (Auster & MacRone, 1994). The primary instructor behaviors and perspectives believed to affect class participation include the following: 1) effectively communicating to students that participation is valuable to all and highly desirable; 2) providing a safe environment in which differing views are acceptable; and 3) expressing that student's unique experience and/or knowledge portrayed in one's comments benefits the whole class (Howard et al., 2002). Auster and MacRone reported that students are more likely to participate when they are encouraged to make voluntary comments, addressed by name, provided time to formulate an answer, given positive feedback, and asked higher-order rather than factual questions.

Although higher-order questions are thought to be highly desirable in achieving educational goals and in creating enriched class discussions, Barnes (1983) discovered that a mere 18% of instructors' questions are above the lowest cognitive level of Bloom's (1956) taxonomy. Bloom's (1956) typology of learning, often cited in articles pertaining to participation and/or critical thinking, is divided into six categories, with the lowest being the simplest and most concrete: 1) knowledge, 2) comprehension, 3) application, 4) analysis, 5) synthesis, and 6) evaluation. In keeping with the taxonomy, instructors must bear in mind that the more complex the question, the more time may be required for students to formulate a well-thought-out response.

Thus, it appears that methods of increasing student participation in class discussion are worth researching. Often, instructors either use an "open discussion" or "cold-calling" format. The former reflects an arrangement in which students are able to comment in a free-flow discussion and the latter is characterized by students being singled out for a response by the instructor. According to Bean and Peterson (1998), instructors using the open discussion format are more likely to welcome all comments and instructors using the cold-calling approach are more likely attempting to assess the amount of knowledge and understanding students possesses. Each approach has benefits and limitations. Open-discussions provide an environment in which students can formulate their ideas, participate when they feel comfortable, and speak when they have something beneficial to add to the discussion. However, the cold-calling approach is more likely to force students to remain focused on course material and class discussion in order to respond intelligently when questioned. Additionally, in the cold-calling format instructors may either randomly or systematically call upon students in an effort to balance discussion among students.

Other variables affecting class participation pertain to the environment. For example, the optimal classroom seating arrangement is thought to be one in which chairs are arranged in a circle (Gladding, 2008). Howard and Henney (1998) found that student participation also is predicted by course size, with smaller classrooms generally yielding greater participation. Montello's (1988) literature review found no significant relationships between where students sit in the classroom and their achievement. However, those sitting in the front and center of the classroom were more likely to participate than those sitting elsewhere. In addition, data indicate instructors' attempts to rouse student participation wane throughout the semester. Evening

classes were positively related to student-initiated discussions, but amount of discussion and direct questions initiated by the instructor were not significantly related to time of day (Howard & Henney).

Grading Student Participation

Jacobs and Chase (1992) have been critical of grading participation because they suggested personality to be the primary variable being graded. Therefore, it can be argued that introverted and/or shy students would be unjustly punished under a credit for participation arrangement. Also, Jacob and Chase argued that recording participation is extremely difficult and hard to rationalize when students complain about their final grade, particularly when instructors assign participation grades retrospectively. This is especially problematic when participation credit is used to help or punish students on the cusp of a desired letter grade. Additionally, Davis (1993) recommended that student participation not be graded due to its highly subjective nature. One solution suggested by Howard and Henney (1998) is for students to grade themselves at the conclusion of each class. A less subjective participation measure was developed by Foster et al. (2009), who required students to self-record their comments immediately after being made. This method helped minimize criticism of comment record keeping and credit assignment. Also, this system was designed to more equally distribute comments among students and encourage participation from those initially disinclined to comment (Foster et al.). Having a blank card in front of a reticent student likely serves as a reminder that he/she has not participated and a card with several comments cues a verbose student that he/she may be dominating the discussion (Krohn, 2010; Krohn et al., 2010).

As an alternative to grading participation, Hodge and Nelson (1991) incorporated differential reinforcement principles to distribute discussion and facilitate participation from less talkative students. As the students made comments, the instructor placed a mark below their names displayed on the chalkboard. A person could either receive a plus mark for increasing his/her participation or for decreasing his/her participation. Students judged to be "over-participants" only received marks when they did not initiate participation, did not interrupt, or responded when addressed. Reticent students were reinforced merely for attempting to participate (any type of comment, hand raise, eye contact). Moderate participants received a variable ratio schedule of reinforcement. This procedure improved distribution of participation across students. One drawback noted by the authors was that some students were upset that they did not receive marks (reinforcement) for their comments. It was noted that these students typically commented frequently. Given that only 14 students were in the class, another limitation is the impracticality of such a system in a larger class.

Even though students acknowledge both an increase in quality and quantity of class discussions stemming from course credit for participation, they remain ambivalent toward such an arrangement (Sommer & Sommer, 2007). Sommer and Sommer also suggested one of the reasons students give for not desiring a participation grade is that students will make tangentially related comments just to meet the "quota." In addition, students most enthusiastic about a participation grade are generally already involved in discussion. Regardless of drawbacks to crediting participation, many other authors maintain that doing so is of great importance (Bean & Peterson, 1998; Dallimore, Hertenstein, & Platt, 2004; Howard & Henney, 1998; Mainkar,

2008). Also, Janzow and Eison (1990) suggested that if class participation is not awarded credit, then students will fail to be engaged and not view discussion as a worthwhile endeavor.

Boniecki and Moore (2003) recommended token economies incorporating backup reinforcers to be an effective method for increasing student participation. Tokens can be exchanged for extra credit or for dropping a low grade on a homework, quiz, or exam. Boniecki and Moore also reported that students required less time to formulate a thought under the token economy. They reported that on average students only took 1 second under the token economy and 6 seconds in baseline to respond to an instructor's question. Although a token system may eliminate "dead time," it may also create an atmosphere of impulsive responding in which students have not developed a well-organized response.

Another credit issue is whether participation should be awarded with regular course credit or with extra credit. No studies were identified that compared the two arrangements; however, Boniecki and Moore (2003) found that an extra credit arrangement can effectively increase participation levels, even in a large class. As with awarding no credit, awarding extra credit for participation may devalue the behavior. Also, many students may not find extra credit a sufficiently powerful incentive to participate in discussion. Therefore, the proper amount of credit to be awarded, the type of credit (i.e., regular or extra), and the optimal time to award credit for participation are important factors. Foster et al. (2009) provided two points each day for participation during one semester and three points per day the following semester. Results revealed low-responders were significantly more likely to participate in credit than non-credit units. Thus, the arrangements were effective in promoting participation by initially reticent students.

Assessing Student Participation

Participation could be evaluated in a variety of ways: end-of-course surveys, peer-recording and/or rating, and self-recording. Although an end-of-course survey may be the most straightforward and non-invasive method, it suffers from a high degree of subjectivity and may increase grade challenges from students. Further, students tend to inflate participation levels almost twice as much as were observed (Howard et al., 2002) and rate themselves as more involved in discussion than do their peers or instructor (Burchfield & Sappington, 1999).

Melvin (1988) suggested peer recording of participation to have sufficient inter-rater agreement and validity. Yet, Love (1981) did not find peer ratings to be viewed as acceptable by students. Mainkar (2008) also voiced criticism for such methods. He suggested that a daily peer ranking is impractical and using peer nominations to rank the students who participate the most (top 20%) is insufficient for the remainder of the class. He also warned that student absences and recency effects would be difficult to overcome. Furthermore, daily peer ratings may be distracting, lead to students bribing each other, and reduce valuable class time.

Results of a meta-analysis examining the reliability of various academic self-assessments found that students do not under-rate or over-rate themselves consistently; the higher the level of course the greater the accuracy; and greater likelihood of poor students' inflating their participation scores and better students' under-rating their scores (Falchikov & Boud, 1989). However, this meta-analysis targeted overall self-assessment rather than participation specifically. Under credit producing contingencies, Krohn et al. (in press) found an 84% agreement between observers and students who recorded their comments on specially designed record cards. Under non-credit contingencies, the inter-rater agreement for specially designed

record cards fell to 75% agreement with observer's records. However, it is important to note that the study found most students to be under-reporting rather than over-reporting. Krohn et al. claimed that frequent and objective self-recording measures are superior to general ratings due to increased accuracy in grading and increased quality (on-topic, non-repetitious) of participation.

Another major issue to be explored is the interaction of quality and quantity of participation when more students participate in class discussion. Although such arrangements have shown quantitative increases in participation (Boniecki & Moore, 2003; Krohn et al., 2010; Krohn et al., in press; Sommer & Sommer, 2007), systematically assessing impact on the quality of participation is more difficult (Mainkar, 2008). Recording the quality of responses requires raters with a thorough understanding of the course material. Whether students are able to accurately rate the quality of their own comments remains virtually unexplored. Krohn (2010) limited recording the quality of participation to recording comments as either relevant or non-relevant. Results revealed few non-relevant self-recorded comments. Participation contingencies produced no trends in quality of comments. Student and observer agreement ranged from 0 to 30%. However, agreement between observers was very high. Differences between student-observer and inter-observer ratings suggested a difference in interpretation of the operational definition of quality. Using a similar arrangement, Aspiranti (2010) limited students self-recording of comments to either timely or repetitious. Similar to Krohn's study, few repetitious comments were self-recorded and no contingency increased repetitious comments. Thus, claims of students making irrelevant comments to gain participation credit appear unsubstantiated (Gilson, 1994; Mainkar, 2008).

Participation History and Beliefs

Krohn (2010) and Aspiranti (2010) used a 50-item participation survey to examine its predictive potential of low-, medium-, and high-participants. Both researchers conducted several principal components factor analyses with varimax rotation. Krohn found three factors: 1) Expectation for Discussion in College Classes, 2) Personal Benefits of Participation, and 3) Personal History and Confidence Regarding Participation. Analysis of variance results found low-, medium-, and high-participant groups to differ significantly on the total survey, as did pairwise comparisons of the groups. The only exception was the Personal Benefits of Participation factor on which only the low- and high-participant groups were significantly different. The greatest differences in mean factor scores were for the personal History and Confidence Regarding Participation, followed by Expectation for Discussion in College Classes, and Personal Benefits of Participation, respectively. A Discriminate Function Analysis revealed the survey to correctly predict students' participation level per unit 68.1 to 77.6% of the time.

After conducting principal component analysis with varimax roation Aspiranti (2010) found seven factors: 1) Personal History and Preference regarding Class Participation, 2) Impact of Discussion on Course Value and Grades, 3) Cognitive and Affective Investment in Class Discussion, 4) Relevance of Discussion, 5) Possible Impediments to Discussion, 6) Responsibility for Discussion, and 7) High Quality Contribution to Discussion. Aspiranti reported significant differences between low-, medium-, and high-participant groups on the total survey. Using a Tukey post hoc test, Aspiranti found significant differences on mean test scores between low- and high-responders, but not medium- and high-responders. Significant differences between all three participation groups were found on Personal History and Preference

regarding Class Participation, but only significant differences between low- and high-participant groups on Impact of Discussion on Course Value and Grades as well as Possible Impediments to Discussion. Aspiranti's best factor predictor was History and Preference (items 1-4, 9, 16, 33, and 38), which is similar in name and item composition to Krohn's (2010) best factor predictor (Personal History and Confidence Regarding Participation).

Critical Thinking and Participation

One of the primary goals of encouraging class discussion is to promote critical thinking skills. Smith (1977) administered the Watson-Glaser Critical Thinking Appraisal (WGCTA) in 12 college courses and found significant positive relationships between critical thinking and participation, encouragement, and peer-to-peer interaction. Also, Garside (1996) showed that the traditional lecture method was inferior to group discussion in assessing student learning. However, it is important to note that the two groups did not differ in terms of promoting critical thinking. McCleary, Foster, and Williams (2011) found that students who participated frequently in large college courses had higher critical thinking scores than those who participated less frequently. However, this relation does not clarify whether higher critical thinking produces more participation or vice versa.

Past research in the same course used in the current study showed that by the end of the course students high in critical thinking were more likely to accurately evaluate course topics and cite professional sources for their evaluations than those low in critical thinking (Williams, Bain, Stockdale, Isaacs, & Williams, 2003). In addition, several studies have documented critical thinking to be significantly and positively related to multiple-choice exam scores in the course (Wallace & Williams, 2003; Williams, Oliver, Allin, Winn, & Booher, 2003a, 2003b; Williams,

Oliver, & Stockdale, 2004; Williams & Worth, 2002). Also, Williams et al. (2003b) found that students were more likely to make knowledge gains throughout the course than critical thinking gains.

Despite the positive relationship between critical thinking and exam performance, Williams and Stockdale (2003) reported that low critical thinkers can still perform well in a course. Williams and Stockdale first characterized students as low-performing low critical thinkers (LPLCT), high-performing low critical thinkers (HPLCT), and high-performing high critical thinkers (HPHCT). The HPHCT students outperformed all other groups on practice and unit exams, ACT scores, and GPA. However, HPLCT students obtained the same attendance record and earned only slightly lower GPAs than HPHCT students. In addition, HPLCT students were better at notetaking, and improved critical thinking scores more throughout the course than the comparison groups.

Framework for the Current Study

The current study can be considered a follow-up to Aspiranti (2010), Foster et al. (2009), Krohn (2010), Krohn et al. (2010), and Krohn et al. (in press). We used similar research procedures, while extending the research literature concerning student participation. Previous research showed high reliability in student self-recording procedures and minimal reactivity of this procedure on student participation. Specially designed student record cards have proven to have adequate submission rates and are instructor friendly in that grades can be assigned quickly (Foster et al., 2009).

The current study addresses a major limitation of the Aspiranti (2010) study, which found a weak treatment effect when specific days to credit participation were randomly selected at the

end of the semester. Aspiranti's study was a follow-up to the studies of Foster et al. (2009), Krohn (2010), Krohn et al. (2010), and Krohn et al. (in press), all of which gave participation credit immediately after completion of selected units. The aforementioned studies documented the reliability of students self-recording comments, the minimal effect of self-recording on participation, and the effectiveness of immediate crediting of participation, on increasing the percentage of initially reticent students. Increasing student participation throughout the course, while decreasing the amount of instructor rating time was the primary goal of Aspiranti's study. Aspiranti randomly selected days in specific units to credit participation at the end of the course. In one section, students were told at the start of Unit 2 that 2 days from both Units 2 and 4 would be credit-producing. The 2 days within each unit were randomly selected by the students at the conclusion of the semester. In a second section, the instructor selected two units at the end of the semester for possible credit and then had students randomly select 2 days for credit within each of those units. A third section randomly selected two units for credit and then 2 days within the selected units for participation credit at the end of the semester. The intent of waiting until the end of the semester to make random selections was to increase the stability of student responding throughout the semester. It was hypothesized that arrangements two and three would produce more stable levels of participation across the semester than arrangement one.

Aspiranti's (2010) results revealed that the contingency in the first section produced elevated participation levels in the known credit units and decreased participation in non-credit units. The other two sections showed minimal participation differences across credit and non-credit producing units. The second section (instructor selected two units at end of unit) experienced a slight decrease in student participation during the course, while the third section

(student selected units and days at end of course) experienced a slight increase. The specific credit contingencies produced no consistent pattern for the latter two sections, but there were participation differences from the start of the course to the end of the course. The second section experienced an overall decrease in participation. Percentage of non-participating students increased from 27 to 47% and percentage of frequently participating students decreased from 21 to 9%. The third section sustained the number of students not participating, and elevated the percentage of students frequently participating from 17 to 22%. Overall, percentage of students participating at different levels when both credit units and credit days within units were designated at the end of the course approximated the probability of credit for participation on those days.

Compared to the results of the Krohn et al. (2010) study, Aspiranti (2010) produced less than desirable results (i.e., a thin random reinforcement schedule selected by students at the end of a course had reduced effects on participation). Thus, the current study seeks a more conservative approach than the Aspiranti study in increasing participation levels, whilst decreasing instructor participation grading time. By making the reinforcer (i.e., credit for participation) closer in proximity (i.e., end of unit) to the student behavior, participation levels should be heightened. Informing students which units, but not which days within the units when participation credit would be awarded, should increase participation levels throughout the selected unit. The random selection of days within units should come at the end of each announced unit in order to save the instructor time compared to recording participation each day within units.

Research Questions

The primary research question addresses the relation of credit contingencies and students' self-recorded participation. Secondary questions target the predictability of participation levels based on survey reports of behavior in prior classes, attitudes toward participation, and degree of critical thinking.

Primary Research Question. The primary research question is whether students' randomly selecting days at the end of previously announced units to credit class discussion promotes participation levels equivalent to those found when students receive credit every day and higher than those found when credit days are randomly selected at the end of the semester (as previously assessed by Krohn, 2010 and Aspiranti, 2010, respectively). This primary research question is an extension of previous research conducted within the Educational Psychology 210 research team in a sequence of three semesters regarding the effects of credit contingencies upon participation levels.

Secondary Research Questions. The secondary research questions target the potential of students' self-reported history of participation in previous classes, attitudes toward student participation, and critical thinking to predict participation. Students scoring high on self-reported history of participation, positive participation attitudes, and high on critical thinking were expected to be more active in course discussion than students scoring low on these variables.

Chapter II

Method

Participants

During the fall semester of 2009 students in three large sections (n = approximately 54 students in each section; total n = 161) of an undergraduate educational psychology course participated in the study. Participants were predominately female (78%), and either sophomores (45%) or juniors (35%). Mean GPA for participants was 3.28. The Teacher Preparation Program requires this course for all students applying to that program. All participants were attending a state university located in the Southeastern United States.

Environmental Context for the Study

The course chosen for the study focused on issues in human development. The course was divided into five units: Physical Development, Cognitive Development, Social Development, Psychological Development, and Values Development. Most units spanned seven class periods and followed a similar format. In the first class period a video was shown; in the next four class periods assigned reading materials were discussed; in the sixth class period a practice exam was scored, a quiz administered, and assigned articles were discussed; and in the final class period students completed a unit exam. The current study focused on the discussion class periods, days 2 through 5. During the third class discussion period of each unit, students' participation was externally monitored by former and current graduate teaching assistants in the course.

Each section met successively for 50 minutes in the mornings of Monday, Wednesday, and Friday with its respective instructor. Although each section had a different lead instructor,

the instructors were enrolled in the same doctoral program (i.e., school psychology) and trained by the same supervising instructor. The prior year, lead instructors served as graduate teaching assistants in the course and occasionally led the course discussion under supervision. Each course followed the same syllabus (with exception of participation contingencies), course materials/content, and course organization.

The course was designed to facilitate class discussion, rather than class lecture, by maximizing opportunities for student participation, engagement, and learning. To maximize student preparation for class discussion, the instructors provided students with printed course notes, PowerPoint slides, and related articles at the start of the course. Students were then directed by the instructor and the syllabus to come prepared and ready to discuss the topics to be covered each day. During course discussion days the instructor presented the class with questions to probe student understanding and comprehension. In addition, the instructor posed questions requiring students to synthesize course material, compare and contrast course concepts, and apply course knowledge to current and events. In an attempt to increase regularity of attendance, instructors also awarded credit for attendance (Mainkar, 2008). Regular attendance was viewed as being beneficial for students to build rapport with one another and, consequently, be able to discuss more controversial topics in a "safe" environment. Also, class discussion was viewed as a modeling opportunity for less conversive students.

In addition, homework was assigned nearly every day (2 through 6). Homework questions covered the material found in the course notes, slides, and articles. Many of the questions prompted students to compare and contrast, synthesize information, and/or draw conclusions based on the information presented. Comprehension question were posted on the

course website as an additional way of engaging students critical thinking abilities. The comprehension questions were provided to students at least the day before the start of each unit. Posting of the questions gave students time to ponder the comprehension questions, perhaps locate additional outside sources, and formulate a response and/or question which he/she could then contribute to the class discussion.

Recording Procedures

Students were required to purchase participation record cards that were specifically designed for this course (see Figure 1). The students were instructed to record their own voluntary comments on these record cards during class days 2 through 5 in each unit. The record cards were designed according to the recommendations of Krohn et al. (in press). For example, Krohn et al. recommended the record cards specify other credit-producing activities (i.e., attendance, proper display of name card, completion of both homework assignments). As such, on days 2 through 5, students could earn up to four points each day for indicating on the record card their engagement in other credit-producing activities even without having commented in class discussion.

Recording student comments. In each unit, voluntary comments and relevance of each comment was recorded by each student on the record card. Students were instructed to write down the main idea or gist of each voluntary contribution to the class discussion. In this manner, the number of comments and accuracy of what one said would be indicated by circling a "T" for Timely or an "R" for Repetitious next to each comment. A voluntary comment could be in the form of a question, response/answer to a question posed by the instructor, or a response/answer to a question posed by another student, without being called upon. The syllabus and the lead

instructors informed students that voluntary comments that led to an interchange on the same topic between instructor and student or student and student(s) only counted as one comment. In addition, if students responded to a question in unison, they were not to count the remark as a voluntary comment. Only individual comments were to be recorded on the record cards.

Students numbered their first three voluntary comments on the front of the record card and wrote any additional comments on the back of the card. Next to each comment were boxes where students marked if the voluntary comment was timely or repetitious. A timely comment was on-topic and non-redundant. Repetitious comments were questions already posed by another student (i.e., all comments not meeting the timely criterion). Students were informed that instructor feedback to their comment would aid in judging repetitious comments. For example, an instructor's response of "that's the same question Jane asked" or "that's a point we discussed a few moments ago" would cue the student that his/her comment was repetitious and the student should closely track the class discussion. This information was contained in the syllabus and frequently reiterated by the lead instructors.

Regardless of the credit contingency in the various sections and units, students were to self-record their voluntary comments on the record cards on days 2 through 5 in each unit. At the end of each class session, students turned in their cards to the instructor. At the beginning of each class period, instructors reminded the students as to the credit contingency for that unit. In addition, this information was displayed on the course website and on the PowerPoint slide shown at the beginning of each class session, with the exception of the baseline unit.

Given that the self-recording procedure produced high reliability in the past (Aspiranti, 2010; Krohn, 2010), the current study only assessed inter-rater reliability one day per unit.

Having observers only one day a unit allowed for a less intrusive classroom environment that more closely approximated a typical classroom. Specifically, Krohn et al. found inter-rater agreement between student and observers to range between 82 and 98% for credit unit and between 75 and 96% for non-credit units. Inter-rater agreement between observers ranged from 93 to 100%. Aspiranti obtained inter-rater agreement between student and the primary observer to range from 81 to 98%. Inter-rater agreement between observers ranged from 88 to 96%. Inter-rater agreement was conducted on the fourth class discussion day in each unit in the previous two studies. In the current study, two observers rated the quantity and timeliness of each student comment and two assessed variability in instructor behavior on the third day in each unit. Observers sat in the front of the classroom where they could to see the students' name cards and accurately record who was making voluntary comments.

The observers were trained in the rating schemes and use of forms prior to the first designated inter-rater day. In an effort to control for instructor differences and to enhance class discussion, instructors provided all students with the same list of potential discussion questions before the first day of class discussion in each unit. Thus, students could formulate well-developed responses to the questions and consider other relevant questions/issues to present to the class. In this manner, even the most reticent students would be able to prepare thoroughly for class discussion. The observers recorded students' voluntary comments and the timeliness of each comment on the Student Discussion Form developed by Krohn (2010) on discussion days throughout the course (see Appendix C).

Recording instructor behaviors. Teacher behaviors were monitored to control potential instructor variability between sections and across units. On the selected day in each unit instructor questions and responses to student comments were recorded by two graduate teaching assistants on the Instructor Discussion Form developed by Krohn (2010) (see Appendix D). Each instructor question was rated as factual or comprehension in nature. Factual questions were defined as questions that could be answered in a straightforward manner from information in the instructor notes. On Bloom's (1956) taxonomy these types of questions would be considered bottom-level knowledge (e.g., recall of information; knowledge of dates, events, places; and knowledge of major ideas). Comprehension questions were defined as those designed to elicit student interpretation, prediction, application, illustration, and compare/contrast responses. As such, a comprehension question would cover the higher levels in Bloom's taxonomy (e.g., comprehension, application, analysis, synthesis, and evaluation).

Instructor feedback was assessed as either positive or negative. Positive feedback consisted of the instructor's restating the main idea of the student's comment or affirming statements as to their accuracy or relevance. Negative feedback was recorded when an instructor provided the student with verbal feedback indicating that his/her comment was either off-topic, repetitious, or incorrect. Also, an instructor's non-response to a student's comment was considered negative feedback.

Crediting Procedures

Students were awarded credit for participation in class discussion on four randomly selected class discussion days across two researcher-selected and pre-announced units. Following baseline, each section alternated between credit units and non-credit units. However,

Section B began with a non-credit unit whereas Sections A and C began with credit units. Two days for awarding participation were randomly selected by students at the end of each credit unit (see Table 1). In each section the first unit was used as a baseline. In Sections A and C, Units 3 and 5 provided no credit for participation. At the end of Units 2 and 4, two class discussion days were randomly selected by students for participation credit. In Section B, Units 2 and 4 were non-credit units, whereas two class discussion days were randomly selected for credit at the end of both Units 3 and 5. For credit-contingent units, students were awarded three points for their first voluntary comment and two points for their second voluntary comment on each discussion day. Therefore, during credit-contingent units, students could earn up to 10 participation points per unit. In addition, at the conclusion of the course students could earn up to 5 or 10 additional points based on their level of participation across the randomly selected days. For example, if a student participated at least once on all of the randomly selected days, the student would receive 5 points; however, if the student participated at least twice on all of the selected days then the student would receive 10 points. Thus, credit for participation was awarded on a relative immediate basis and also on a delayed basis at the end of the course. At the end of the course, 5 points were awarded to students contributing at least one comment on each of the four selected participation days, and 10 points were awarded to students contributing at least two comments on each selected day. Thus, 30 points were available for participating in class discussion, approximately 5% of course credit.

Survey of Student Participation Perspectives

A student participation survey (adopted from Aprianti, 2010 and Krohn, 2010) was posted on the course website. Students earned five points for completing and turning in their

responses to the instructor at the beginning of the course. Students indicated their answers on computer scan forms. The scan forms allowed for electronic entry of student's responses into the database, minimizing error in data entry.

Survey data were used to predict class participation. Also, the information provided a type of learning history from which to speculate why certain students contributed minimally to class discussion, regardless of the credit contingency in place. Questions on the survey were designed to evaluate history of participation, comfort in participation, perceived worth of class discussion in a course, and whether the student or instructor was perceived as responsible for class discussion. Krohn (2010) completed a series of factor analyses resulting in three factors: 1) expectation for discussion in college classes, 2) personal benefits of participation, and 3) personal history and confidence regarding participation. Each factor consisted of seven to eight questions, and all factors were included in the prediction of participation under credit vs. non-credit conditions. Given the internal consistency and predictive validity of Krohn's factors, the same factor structure was used in the current study.

Critical Thinking and Participation

The Watson-Glaser Critical Thinking Appraisal-Form S (WGCTA; Watson & Glaser, 1994) was administered in class on the second day of class as a predictor of participation. The Form S contains 40 items and is the short form of the WGCTA-Form A. The WGCTA is designed to measure critical thinking in adults. The WGCTA Form S reports an internal consistency and test reliability of .81. Form S reportedly correlated with reading achievement measures in the mid .60 range. Correlations between ACT, SAT, and the Form S ranged from the mid .30s to mid .80s (Watson & Glaser). Previous studies examining the predictive validity

of the WGCTA-S have found high critical thinkers to more accurately evaluate course topics than low-critical thinkers (Williams, Bain, et al., 2003); critical thinking to be positively and significantly related to exam scores (Wallace & Williams, 2003; Williams, Oliver, et al., 2003a, 2003b; Williams et al., 2004; Williams & Worth, 2002), and low-responders to have significantly lower critical thinking scores than high-responders, regardless of academic year and gender (McCleary, Foster, & Williams, 2011).

Participation Contingencies and Research Design

Each section contained both credit and non-credit units. Sections were counterbalanced with respect to the sequence of credit and non-credit units. Each section had a baseline unit in which no credit was given for participation and then two credit units separated by non-credit units. Students self-recorded in each unit for class discussion days (2 through 5). Table 1 shows the treatment conditions assigned to each section. Credit versus non-credit participation data were analyzed via criterion-by-inspection and proportion testing.

Chapter III

Results

Information in this chapter is presented in the following sequence: 1) demographic and correlational descriptive data; 2) reliability of student self-recording and inter-rater agreement; 3) inter-rater reliability data of instructor questions and feedback; 4) graphic presentation of student participation levels (dependent variable) under different credit conditions; 5) student participation levels analyzed with proportions testing to determine significance levels of treatment differences; 6) intra-student comparisons of initially low-, medium-, and high-participating students across units and treatment conditions; 7) instructor behavior and potential confounds or interactions with treatment effects; 8) analysis of Participation Survey (e.g., predictive potential of factors); and 9) relationship between WGCTA-S scores and levels of participation (e.g., predictive potential using logistic regression). Data were evaluated primarily on a group basis. All sections were compared under credit versus non-credit conditions. Some small group comparisons were conducted to assess differences between high-participating and low-participating students.

Students' self-reported participation data were compared to raters' participation data for one day each unit. This procedure was conducted in order to identify students who consistently chose not to record their class participation accurately. Two participants were removed from all analyses involving self-reported participation, because they were identified by observers as consistently under-reporting the number of comments made during participation. One student's self-recorded comments totaled across all five units was as much as 24 comments less than

recorded by both independent raters, and the second student self-reported 30 fewer comments than recorded by both independent raters.

Descriptive Demographic Data and Correlations

An ancillary issue was the different levels of participation among students based on their gender and academic classification (e.g., academic year). A point-biserial correlation between mean participation for each student across units and gender (.11) and a Pearson correlation between mean participation for each student and GPA (.04) were not significant. However, a significant relationship ($r = .14$, $p < .05$) was obtained between academic year and mean participation. Table 2 displays the mean level of student daily participation by gender and academic classification. Although females participated in class discussion more frequently than males, this gender difference was not significant. During non-credit participation units, the only significant difference among participation rates and academic classifications were between graduate students ($M = 1.93$, $SD = 1.23$) and sophomore students ($M = .95$, $SD = .94$). Due to the significant difference between graduate students' and sophomore students' participation in baseline, the graduate students were removed from any analyses involving self-recorded participation. Given the small number of graduate students ($n = 8$), they were removed from the data set instead of sophomores ($n = 71$). Freshman ($n = 2$) students also comprised a small sample size. No significant differences were found among academic classifications during credit units.

Reliability Assessments

Data were collected from students and a team of four external observers. Inter-rater reliability of teacher behaviors were assessed by two observers; an additional two observers

assessed student behaviors one day per unit. Records of students and the observers were compared for level of agreement.

Student participation. Students self-recorded their comments on discussion days (2 through 5) in each unit. Student self-recording and observer recording were compared on the third discussion day of each unit (25% of each unit). Mean percent of inter-observer agreements between students and observers and between observers are presented in Table 3. The mean percent of inter-observer agreement between students and observer 1 ranged across units from 62 to 92%, with a mean of 81%. Students and observer 2 agreement ranged from 64 to 92%, with a mean of 80%. Close scrutiny of student-observer agreement revealed a pattern of heightened student-observer agreement during treatment units.

Table 4 shows the means and standard deviations of student participation during inter-rater check days in each unit. Visual inspection of the data did not indicate any consistent pattern of over-reporting across units. However, visual inspection did reveal a general pattern of students' under-reporting participation across units. Students' under-reporting has been found by previous researchers as well (Aspiranti, 2010; Krohn, 2010; Krohn et al., 2010; Krohn et al., in press).

Student under-reporting. Appendix E displays detailed information regarding students' proclivity to under- and over-report. Student and observer agreement and disagreement (less than or greater than) is listed for each inter-rater reliability day. A proportion-testing procedure explained by Ferguson and Takane (1989, pp. 198-200) was used on a TI83 graphing calculator to determine the significance of the difference in proportions of students' under- and over-

reporting participation. In non-credit units, between 11.18 and 14.29% of the students under-reported their participation during inter-rater reliability days. The average percent of students under-reporting during non-credit units was 13.91 (21/151; where 21 is average number of students under-reporting in both non-credit units and 151 is the number of students in the study). In credit units, between 4.97 and 7.45% of the students under-reported their level of participation. The average percent of students under-reporting during credit units was 6.21 (10/151). The proportion of under-reported cases in non-credit units was significantly more than in credit units (p = .019). As previously found (Krohn, 2010), students generally under-reported when they had already recorded enough instances to maximize participation credit. There were only 11 instances on inter-rater days in which a student under-reported to the extent that the student would not receive course credit had that day been randomly selected for credit.

Student over-reporting. The average percent of students over-reporting during non-credit units was 3.31 (5/151), whereas the average percent of students over-reporting during credit units was 11.26 (17/151). The proportion of students over-reporting during credit units was significantly more than in non-credit units, p = .004. In contrast, 9.94 to 10.56% of the students across sections over-reported their participation during credit units. Nearly half (20) of the students (43) over-reporting their participation on inter-rater days would have unjustly received full participation credit, had that day been randomly selected for credit. An average percent of 0.66 (1/151) students over-reported beyond credit-level (i.e., observers and student agreed that full credit had been earned, yet the student continued to report comments the observers did not) during non-credit units and an average of 3.97% (6/151) students over-

reported during credit units. The proportion of students over-reporting during non-credit units was significantly less than students over-reporting during credit units, p =.028.

Qualitative ratings. In a similar fashion, students and observers recorded students' comments as timely or repetitious. Table 5 shows the mean percent of inter-observer agreement of timely comments between students and observers 1 and 2 and between observer 1 and 2. Agreement between students and observers ranged from 59 to 92% on inter-rater check days. Agreement of timely comments between observers ranged from 73 to 100%. Table 6 displays percent inter-observer agreement on repetitious comments between students and observers on inter-rater check days. Agreement between observers and students and between observers both ranged from 0 to 100%, across units.

Instructor behavior. Instructor data also were obtained one day per unit. This information was gathered to account for possible differences in instructional behaviors across instructors and units within sections. Instructors were compared on the type of questions (comprehension or factual), number of questions asked, and type of feedback provided (positive or negative). On the third day of each unit, inter-observer agreement between observers was assessed on these dimensions. Data were calculated by dividing the higher frequency count by the lower frequency count for each type of question posed. Percent of agreement for factual and comprehension questions ranged from 67 to 100%, across units (see Table 7). Agreement did not differ by contingency or unit. However, percent agreement was generally lower for factual questions than for comprehension questions. The number of factual questions posed ranged from four to 20 per class, whereas the number of comprehension questions ranged from 15 to 34. Thus, disagreements among observers were greater for factual questions than for comprehension

questions. Table 8 displays the mean percent of inter-observer agreement between observers for instructor feedback (positive and negative) to students on the inter-rater check days in each unit. Mean percent of inter-observer agreement ranged from 83 to 96% on positive feedback and from 0 to 50% on negative feedback. Due to considerably fewer recorded instances of negative feedback than positive feedback, a small number of disagreements between observers more greatly affected the percent of agreement for negative feedback than for positive feedback. Percent of agreement between observers ranged from 76 to 100% for the total number of questions asked.

Effect of Credit Treatment on Levels of Student Participation

The effect of credit contingencies on percentage of students participating at different levels across treatment and non-treatment phases in each section was assessed by visual analysis and via proportions testing. Number of comments from each student was converted to show the percent of students participating at various levels. In order to determine the equivalence of students' initial course knowledge and participation levels, a one-way analysis of variance (ANOVA) was performed on Unit 1 exam and participation data across sections. This procedure provided an additional check on whether treatment effect might be related to initial differences in course knowledge and participation levels.

Equivalence of knowledge across sections. Mean differences in Unit 1 exam scores across sections might have indicated a difference in psychoeducational knowledge and/or academic ability. Theoretically, these section differences might have impacted the frequency of participation. Therefore, a one-way ANOVA was conducted on Unit 1 exam scores across sections to determine if significant differences in mean Unit 1 scores existed during baseline.

Across sections the Unit 1 exam mean scores were virtually identical (Section A = 38.96, Section B = 39.48, Section C = 38.87 out of a possible 50). The ANOVA showed no significant differences across sections, $F(2, 157) = .22$, $p = .806$. Examination of the results shows an equivalence of course knowledge in the initial unit of each section.

Equivalence of baseline participation across sections. An ANOVA was conducted to check whether significant differences in mean levels of participation existed across sections during baseline (Unit 1). No significant differences were found, $F(2, 149) = .25$, $p = .289$. Previous research (Aspiranti, 2010; Foster et al., 2009; Krohn, 2010; Krohn et al., in press) on which the current study was predicated found artificially high baseline levels of participation. The researchers assumed that students' familiarity with the course material increased baseline levels of participation. The current study found the same effect. Section B contained an extended baseline (i.e., Unit 2 was a non-credit contingency). The mean number of student comments per day for Section B dropped from 1.64 in Unit 1 to 1.10 in Unit 2. Section C, which was under a participation credit contingency in Unit 2, decreased from a mean of 1.46 in Unit 1 to 1.35 in Unit 2. However, the mean participation per day increased from Unit 1 (1.35) to Unit 2 (1.44) in Section A, which was also under a participation credit contingency in Unit 2. The lack of increase in participation from Unit 1 to Unit 2 in sections under a credit contingency was found by the previous authors as well. It is worth reiterating that the participation credit contingencies were meant to balance the level of participation across students, not increase overall participation.

Visual analysis of mean participatory levels. Figures 2 through 5 represent the percent of students in each section who participated in class discussion at particular levels under credit and non-credit conditions. The same levels of participation as reported by Aspiranti (2010) and Krohn (2010) were used in the current study. The various levels were based on the number of students commenting each day at the following levels: 1) no participation (0 comments), 2) credit-level participation (1 to 2 comments), 3) frequent participation (3 to 4 comments), and 4) dominant participation (5 or more comments). Visual analysis of the figures reveals greater within group variability in percentage of students at different participation levels during credit than non-credit units. Figure 2 shows a greater percentage of students elected not to participate during non-credit units. Specifically, in Section A the mean percent of non-participants dropped from 62 in non-credit units to 32 in credit units, Section B from 46 to 24, and Section C from 67 to 24. There were no overlapping data points between credit and non-credit units in Sections A and C; however, one non-credit unit data point overlapped with three credit unit data points in Section B.

It is important to note that Section B's baseline was followed by a non-credit unit, in contrast to Sections A and C. When examining Unit 2 across sections, one can see that in Sections A and C, in which the credit contingency was in effect, the percent of non-participants did not change substantially from Unit 1 to Unit 2. However, the percent of non-participants in Unit 2 of Section B, a non-credit unit, did increase substantially from 16 to 43.

Figure 3 displays the percent of credit-level (1 to 2 comments per day) participants. Visual analysis consistently shows the percent of credit-level participants to increase during each credit unit. Across sections, the mean percent of credit-level participants was 29% during non-

credit units and increased to 53% during credit units. There were no overlapping data points between credit and non-credit units in Sections A and C. On the other hand, two non-credit unit data points overlapped with one credit unit data point in Section B. Figure 4 presents the percent of frequent participants (3 to 4 comments per day) each day. Section B experienced the greatest differences in the mean percent of frequent participants across units, 24% in credit units and 16% in non-credit units. Section A had a mean of 12% in non-credit units and 18% in credit units. Section C's mean percent of frequent participants rose from 7% in non-credit units to 13% in credit units. Overall, there was a 6% mean difference percentage of frequent participants between credit and non-credit units. There were multiple overlapping data points in the percentage of frequent participants between credit and non-credit units in all three sections. Figure 5 shows some variability across sections in the percentage of dominant participants (5+ comments per day) under credit and non-credit conditions. The mean number of dominant participants in Section A showed a slight decrease from non-credit to credit units (1.26 to 1.17%), Section B declined from 0.9 to 0%, and Section C increased from 0 to 0.27%. Overall, the percentage of dominant participants was virtually the same decreasing only 0.24% in non-credit to credit units. Several credit and non-credit unit data points overlapped in all three sections.

Proportions testing. Statistical proportions testing (Ferguson & Takane, 1989) was used to quantify changes in the proportion of students participating at various levels under credit and non-credit contingencies. Differences between treatment and non-treatment units were compared via proportion values. Within each participation level, each unit was compared across sections and units (including baseline) within sections.

Across sections, proportion testing of non-participants (0 comment level per day) showed all adjacent treatment pairs to be significantly less ($p = < .05$) in credit units than non-credit units (see Table 9). For example, in Section A the first treatment pair (credit to non-credit) resulted in a significant increase in non-participants ($p = .001$). The second treatment pair in Section A resulted in a significant decrease ($p = .003$) in non-participants from the non-credit to credit phase. In addition, the overall comparisons of credit to non-credit units were significantly different for Section A ($p = .04$) and Section C ($p = .003$), but not for Section B ($p = .12$).

Proportion testing of credit-level participants (1 to 2 comment level) in Section A and C showed all adjacent treatment pairs to be significantly different ($p < .05$). However, in Section B the first ($p = .12$) and second ($p = .07$) treatment pairs did not differ significantly. Overall comparisons between credit and non-credit units revealed Section C ($p = .009$) to be significantly different, but Section A ($p = .07$) and B ($p = .23$) were not significantly different.

Examination of adjacent treatment pairs for frequent participants (3 to 4 comment level) across sections revealed only one significant difference ($p = .03$), which occurred in the second adjacent treatment pair in Section C. Overall comparisons of credit to non-credit units across sections showed no significant differences between credit and non-credit units. Proportions testing of dominant participants (5 or more comments) revealed no significant differences between adjacent credit units or in overall comparisons of credit to non-credit units.

In the main, there were more non-participants during non-credit units than credit units, and fewer credit-level participants during non-credit units than credit units. The number of frequent participants remained fairly consistent during credit and non-credit units. Also, the

number of dominant participants (ranging between 0 and 1) remained extremely low across credit and non-credit units.

Low- and Non-Participants

Low-, medium-, and high-participating students were evaluated using intra-subject comparisons based on baseline designation. In keeping with Krohn's (2010) and Aspiranti's (2010) classification scheme, low-participants were classified as those with an average of 0.5 comments or less per day in Unit 1 (baseline) and constituted the bottom fifth (22%) of the participation distribution. The majority of students initially classified as low-participants remained low-participants throughout the course. Across units, there was a maximum of 85 students who initially met the low-participant criteria. The number of students (85) constitutes 53% of the total participants. Out of the 85 low-participant students 33 (39%) remained low-participant students across at least 4 of the 5 units. In Unit 1 there were 33 low-participant students, with 14 (42%) of those students non-participants (did not participate at all).

Figure 6 displays the percent of initially low-participant students who participated in subsequent units. Across sections there was a consistent increase in the mean percent of participating students during credit units (with baseline included in the mean for non-credit units). Mean percent of initially low-participants who subsequently participated in Section A rose from 12% in non-credit units to 31% in credit units; in Section B it increased from 9 to 18%; and in Section C it rose from 17 to 42%. Proportion testing, using the mean number of low-responding participants who participated in subsequent units divided by the mean number of low-responding participants per unit, resulted in no significant differences ($p = .15$) between credit (4/13) and non-credit (2/14) units in Section A; no significant differences ($p = .50$)

between credit (1/6) and non-credit (1/6) units in Section B; and no significant differences (p = .35) between credit (3/10) and non-credit (2/9) units in Section C. The only significant (p = .03) difference between adjacent treatment and non-treatment unit pairs across sections was between the first credit (5/13) and non-credit (1/13) pair in Section A. Thus, the credit contingency did not appear to have an effect on initially reticent participants.

Table 10 shows the number of low-participants in Unit 1 who fell into low (0 to .50 comments per day), medium (.51 to 1.99 comments per day), and high (2 or more comments per day) categories in subsequent units. As can be seen, the majority of low-participants remained in that category for the subsequent of course units, with little regard to the credit contingency. In Section A, the percent of low-participant students becoming medium- or high participant students was 40% during the initial credit unit and 13% in the second credit unit; 17% (first credit unit) and 17% (second credit unit) in Section B; and 57% (first credit unit) and 43% (first credit unit) in Section C.

In addition, initial non-participants were evaluated regarding their participation in subsequent units. The non-participants (0 comments each day) subgroup of low-participants also increased their participation under credit contingent units. However, they never commented during non-credit units. In credit units, 57% (first credit unit) and 29% (second credit unit) of initially non-participants participated in Section A, 25% (first credit unit) and 50% (second credit unit) in Section B, and 33% (first credit unit) and 33% (second credit unit) in Section C.

Medium- and High-Participants

Again, using the Krohn (2010) and Aspiranti (2010) definitions, the researcher classified students who contributed between .51 and 1.99 comments per day in baseline as medium

participants. Students contributing 2 or more comments in baseline were considered high-participants. Table 11 displays information regarding the medium participants. During credit units, a mean of 21% of Section A's initially medium-participant students moved into the low-participant category and 25% into the high-participant category. In Section B a mean of 20% of the students moved to the low-participant group and 42% into the high-participant group. In Section C 22% moved to the low-participant group and 31% moved to the high-participant group. In non-credit units, the majority of students tended to move from the medium group into the low group. During non-credit units, a mean of 55% of medium-participant students fell into the low-participant group and 10% rose to the high-participant group in Section A, in Section B 67% became low participants and 12% became high participants, and in Section C 63% fell in to the low-participant category and 10% moved to the high-participant category.

In general, the high-participants (2 or more comments per day) were more likely to record fewer (less than 2) comments during non-credit units than during credit units. The mean percent of high-participants moving into either a low- or medium-participant category during credit units was 21% for Section A, 22% for Section B, and 31% for Section C (see Table 12). During non-credit units the mean percent of high participants moving into either a medium- or low-participant category rose to 59% in Sections A and B and 83% in Section C.

Timely and Repetitious Student Comments

Students rated their comments as timely or repetitious on the specially designed record cards. The intent of this arrangement was to help students pay attention to the class discussion and self-monitor the quality of their comments. Located next to the space designated for students to write the gist of their comments was a box for them to mark if their comment was

deemed timely or repetitious (described in the Methods section). In addition, two observers judged students' comments on the designated inter-rater days. Table 13 displays the total number of timely and repetitious comments recorded by the students and observers. The number of repetitious comments reported by students in a unit ranged from 0 to 6 per unit; likewise, the number of repetitious comments recorded by the observers ranged from 0 to 7. In contrast, students' recorded between 11 and 38 timely comments and observers recorded between 11 and 36 per unit.

Visual analysis of the number of students' timely and repetitious comments reveals no consistent trends across credit contingencies. Due to the limited number of reported repetitious comments and the poor agreement between students and observers regarding the occurrence of these comments, no conclusions can be proposed at this time regarding the frequency of repetitious comments under the credit and non-credit contingencies.

Interaction between Treatment and Instructor Behavior

Instructor questions. The number of questions (factual, comprehension, and total) presented by the lead instructor on discussion days were recorded on inter-rater days (see Table 14). In addition, Figure 7 displays the number of total questions per unit posed by each section's instructor on inter-rater days. Visual analysis of Table 14 and Figure 7 reveals no one instructor to consistently ask more questions (total) than another instructor across units. However, the Section A instructor did pose fewer total questions than the other two instructors in three of five units. Figure 8 displays the mean number of comprehension questions posed in each section. Again, no one instructor posed more comprehensions questions than another (across units), but the instructor of Section A did pose the fewest comprehension questions in the first three units.

Examination of the percent of comprehension questions posed per unit in each section shows no one section to pose greater or lesser percent of comprehension questions than another section (see Figure 9). Also, there was no discernable trend among the number of factual questions being posed across sections (see Table 14). Proportions testing revealed no significant differences between the number of comprehension or factual questions posed in credit and non-credit units in Section A and Section C. However, Section B had significantly more comprehension ($p = .001$) and factual questions ($p = .001$) in non-credit units than credit units.

Subsequently, a Chi-square analysis was conducted to determine any significant differences between the three instructors and units in total questions (factual and comprehension) and total comprehension questions asked. The analysis revealed no significant differences in total questions asked across instructors and units, $X^2(8) = 5.43$, $p = .711$. Also, the analysis of total comprehension questions posed across instructors and units revealed no significant differences, $X^2\ (8) = 6.111$, $p = .635$.

Instructor feedback. Type of instructor feedback (positive or negative as defined in the Methods section) was recorded on inter-rater days. Unit means and standard deviations are presented in Table 15. Visual analysis of mean positive feedback per unit shows that 67% of the adjacent treatment pairs were higher in credit units than non-credit units. The proportion of positive feedback to timely responses and the proportion of negative feedback to repetitious comments were all 1.0, indicating appropriate feedback was given according to the type of comment presented.

Figures 10 and 11 display differences within and between sections of the number of positive and negative feedback comments presented to students on inter-rater check days. Visual

examination of both figures shows only minor differences between and across sections and units. A mixed-design ANOVA analysis was performed to determine any significant differences between and across sections and units in the type of feedback (positive and negative) instructors gave. The analysis revealed a main effect for section and no significant main effect for unit or section by unit interaction, F (2, 663) = 3.45, p = .032. A LSD post-hoc analysis revealed Section B (M = 1.50) to be significantly higher in the amount of positive feedback given by instructor to students than Section A (M = 1.18, p = .021) and Section C (M = 1.20, p = .047). Analyses revealed no significant main or interaction effects of the quantity of negative feedback given by instructor to students across sections and units.

Survey Results

Krohn (2010) conducted a principal components analysis with varimax rotation to derive three primary factors: 1) Personal Benefits of Participation (Personal Benefits), 2) Expectation for Discussion in College Classes (Expectation), and 3) Personal History and Confidence Regarding Participation (History/Confidence). The same three factors were used in the current study. Cronbach's alpha was used to calculate internal consistency for all 50 survey items (.90), all 23 of the combined factor items (.91), the 8 item History/Confidence Factor (.86), the 8 item Expectation factor (.82), and the 7 item Personal Benefits factor (.74). All were above Garson's (2008) social-sciences research criterion of .70, meaning that each of the combinations could be considered a scale.

Mean survey factor scores. An ANOVA was conducted to determine significant differences between low-, medium-, and high-participant groups on the Participation Survey. A significant difference among participation levels on the total survey was found, F (2,149) =

14.27, $p < .001$. Tukey post hoc analyses revealed significant differences between low (M = 167.00, SD = 18.42), medium (M = 178.19, SD = 18.75), and high (M = 190.00, SD = 12.57) participation levels. In addition, a mixed ANOVA using the participation groups as the between variable and the three survey factors as the repeated measure resulted in a main effect for the three survey factors, $F(1.70, 243.45) = 11.30$, $p < .01$, a main effect for the participation levels, $F(2, 147) = 21.80$, $p < .001$, and a significant interaction effect for the survey factors and participation groups, $F(3.31, 243.45) = 3.70$, $p = .01$. Thus, there was a significant difference in scores on the participation survey factors depending on students' participation group.

Further analyses (i.e., a series of ANOVA's) of the interaction effect revealed that the low- (M = 21.64), medium- (M = 26.53), and high- (M = 25.99) participation groups were all significantly different on the Personal/History factor. The low group was significantly lower than the medium and high participation group at the $p < .001$ level and the medium group was significantly lower than the high groups at the $p = .011$. On the Expectation factor the low participants (M = 24.39) were significantly lower than the medium participants (M = 27.12, $p < .001$) and the high participants (M = 29.71, $p < .001$), while the medium and high participants were significantly different at the $p = .016$ level. Likewise, examination of the Personal Benefit factor showed the low-participation group (M =23.00) to be significantly different from the medium group (M = 25.36, $p = .009$), and the high group (M = 27.77, $p < .001$) and the medium group to be significantly different from the high group at the $p = .012$ level.

Individual item analyses. Means and standard deviations reveal which survey items students regard as being the most and least important and/or influential. The mean scores for the low-, medium-, and high-participant groups for individual items on the three survey factors and

discarded items are presented in Appendix G. Multivariate analyses (one for each of the three survey factors) were conducted to examine the item-by item scores for all three participation groups. A significant overall effect was obtained for the items on the History/Confidence factor, $F(14, 280) = 3.17$, $p < .001$, the Expectation factor, $F(16, 280) = 3.23$, $p < .001$, and the Personal Benefits factor, $F(16, 276) = 2.26$, $p = .004$. A Bonferroni post-hoc analysis compared item means across all participation groups. Survey item number 22 (effect of personal participation in class discussion on concentration in class) from the History/Confidence factor, item 33 (former teachers characterization of participation) from the Expectation factor, and item 12 (general effect of participation on one's own learning) from the Personal Benefits factor were significantly different across all participation groups.

Items significantly different between the low and high groups and low and medium groups, but not between medium and high groups, were items 23 (History/Confidence), 1, 3, 16, and 38 (Expectation), and 2 (Personal Benefits). The only items yielding significant differences between low and high groups were 25 and 29 (History/Confidence), 4 and 9 (Expectation), 8, 14, 20, and 42 (Personal Benefits). Item 36 (Personal Benefits) was significantly different between only the low and high groups and the medium and high groups. Three items (22, 24, and 44) from History/Confidence, one item (31) from Expectation, and one item (34) from Personal Benefits did not yield any significant differences between participation groups.

Logistic regression. Logistic regression was conducted on the three participation factors and the total survey to determine how well they predicted students to be in low (0 to 2 comments per unit) or high (8 or more comments per unit) participant groups. The distinction between low and high participation groups was made to represent a significant separation between the two

groups. The two groups were designated in the same manner as the low- and high-participation groups used for intra-subject comparisons. These distinctions are consistent with previous research by Aspiranti (2010) and Krohn (2010).

Table 16 shows that the predicted classification of students into low or high participation categories were significant for all units. Chi-square values were significant in each unit and ranged from $p < .05$ to $p < .001$. Cox and Snell R Square values ranged from .04 to .14. Odds ratios ranged from 1.03 to 1.05, meaning students were 3 to 5 % more likely to be classified in the high group (i.e., participate) with every one unit increase in the total participation survey scores. The percent of students classified correctly for all the sections (combined) ranged from 63.6 (Unit 4) to 81.8% (Unit 5). Table 17 displays the percent of cases correctly classified into low and high groups each unit. Examination of cases correctly classified into low and high groups for each unit shows an inconsistent prediction of group membership with the percent of correctly classified cases in the low group ranging between 0 and 100 (mean across units was 78.94) and the high group ranging between 0 and 100 (mean across units was 38.16).

Logistic regression was also performed for all three survey factors for each unit (see Table 18). The full model (all three factors) significantly predicted participation classification as low or high in all units, ranging from $p < .05$ to $p < .001$ and R^2 ranging from .11 (Units 3 and 5) to .24 (Units 2). Comparison of the three-factor model (Table 18) to the total survey (Table 16) shows R^2 values to be higher for the former in all five units. The History/Confidence factor was the only factor significantly ($p < .05$ to $p < .01$) contributing to the predictability of the three-factor model. History/Confidence was significant in Units 1 through 4, but not Unit 5. Odds ratio values for the History/Confidence factor were between 1.10 and 1.27 across units, meaning

students were 10 to 27% more likely to be included in the high participation group with every one unit increase in the three-factor model scores. No other factors in the three-factor model contributed significantly to the combined factor predictive potential. Expectation and Personal Benefits had standardized Beta coefficients close to 0 and odds ratios close to 1, meaning that the two factors did not contribute to the predictability of a student being placed in either the low or high participation groups.

Table 19 displays the percent of cases correctly classified into low and high participation groups based on the three-factor model. The mean percent of correct placement into low-participation group (81.30), high-participation groups (49.22), and overall group classification (74.92) for the three-factor model was better than for the total survey.

Given that History/Confidence was the only factor in the three factor model to significantly contribute to the models' predictive potential, another set of logistic regression analyses were performed that specifically examined this factor's ability to correctly predict a student's participation level (see Table 20). The History/Confidence factor alone produced very similar results to that of the three-factor model. Table 21 displays the percent of cases correctly classified into low and high participation groups by the History/Confidence factor. The History/Confidence factor resulted in a slightly lower ability to predict low participants (77.54%) than the three-factor model (81.30%), virtually the same level of correct placement for the high participants (48.04%) as the three-factor model (49.22%), and slightly lower correct placement overall (72.24%) than the three factor model (74.92%), but still higher than the total-survey model (69.26%).

A potential problem in analyzing logistic regression results is that highly correlated predictor variables may not appear to have predictive potential in the presence of a specific predictor that was strongly correlated with the dependent variable (i.e., colinearity). In the present case the three factors were significantly correlated ($p < .001$). The largest correlation was between Personal Benefit and Expectation ($r = .73$), followed by History/Confidence and Expectation ($r = .72$), and finally Personal Benefit to History/Confidence ($r = .51$). The Expectation factor was least strongly correlated with participation level ($r = .48$, $p < .001$) than were Personal Benefit ($r = .52$, $p < .001$) and History/Confidence ($r = .60$, $p < .001$). Given that Personal Benefit and Expectation were highly correlated with each other and poorly correlated with participation levels, they would not likely decrease or increase the predictive capacity of the survey. However, the History/Confidence factor marginally increased the predictive capacity of the survey.

A closer examination of the History/Confidence factor was undertaken, given that it had the best predictive ability of the three factors in the combined model and was able to predict participation level nearly as well as the three-factor model. Of particular interest was the History/Confidence factor's ability to predict low and high participation under credit and non-credit contingencies (see Table 21). Overall, History/Confidence was able to significantly predict students' participation classification across units ($R^2 = .07$ to .24, all p values $< .001$). History/Confidence was not a better predictor of low or high participants during credit than non-credit units. Logistic regression predicted that there were more low participants during non-credit (188) than credit units (89) across sections. In addition, History/Confidence accurately predicted more high participants in credit (84) than non-credit units (38). Proportions testing

(Ferguson & Takane, 1989) revealed significant differences in the proportions of low participants ($p = .001$) and high participants ($p = .001$).

Critical Thinking Results

Classification of critical thinking scores as low, medium, and high were based on normative percentiles (Watson & Glaser, 1994). The low group consisted of 28% of the sample size and ranged from the 1st to 4th percentile, the medium group constituted 42% of the sample and ranged from the 5th to the 30th percentile, and the high group comprised 30% of the sample and ranged from 35th to the 99th percentile. Proportionality of critical thinking levels was examined to determine differences in participation levels. There were significantly more students in the low critical thinking group who were low participants than high participants, $p < .05$. However, there was no significant difference between the number of students in the high critical thinking group who were low participants than were high participants, $p = .10$.

In keeping with the only available research on participation and critical thinking (McCleary et al., 2011), I placed students' critical thinking scores into low (bottom quartile) and high (top quartile) classifications as described above. Similarly, students were placed into low (bottom quartile) and high (top quartile) groups based on their daily average participation. To be considered a low participant, one's scores must have been below a mean of .57 per day and to be placed in the high participant group one's mean participation must have met or exceeded 1.84. In addition, students' were placed in low ($n = 40$; bottom quartile; exam scores ranged from 0 to 37.05) and high exam-score groups ($n = 42$; top quartile; exam scores met or exceeded a 42.2 exam mean for all five units). For continuity and reasons previously stated, students' in the middle range were excluded.

Logistic regression was used to determine how well students' exam and critical thinking levels (low and high; described above) predicted participation levels (low and high; described above). Results revealed low and high critical thinking and exam scores significantly predicted classification of low and high participation groups, X^2 (2) = 6.30, p = .043. Overall, 70% of the students were correctly classified into low and high participation groups based on their critical thinking and exam scores. Specifically, 66.7% were correctly predicted to be in the low participation group and 75% were accurately predicted to be in the high participation group.

A logistic regression analysis was used to determine how accurately the critical thinking scores predicted participation. Low and high critical thinking groups significantly predicted classification into low and high participation groups, X^2 (1) = 5.60, p = .018. Overall, 67.4% of the students were correctly classified as low or high participants based on their critical thinking scores, 63.6% were correctly classified as low participants, and 70.8% were correctly classified as high participants. Participants were 106% more likely to be a high participant with every one unit increase in critical thinking scores. In other words, students were 2.06 times more likely to be classified as a high participant with every one unit increase in critical thinking scores.

In order to determine if critical thinking scores or exam scores served as a better predictor of participation, the researcher conducted a logistic regression analysis using only exam scores as the predictor. Analysis showed exam scores to be a significant predictor of both low or high participation classification, X^2 (1) = 7.51, p = .006. Overall, 69.4% of students were correctly distinguished as low or high participants based on their exam scores. To be exact, 60.9% of students were correctly predicted into the low participation group and 84.6% were correctly

predicted to be in the high participation group. Thus, exam scores (69.4%) were a slightly better predictor of participation than were critical thinking scores (67.4%).

Chapter IV

Discussion and Conclusion

The current study sought to balance student participation in three large sections (approximately 54 students each) of an undergraduate course by decreasing the number of non-participants, increasing the number of credit-level participants, maintaining the number of frequent participants, and decreasing the number of dominant participants. Krohn (2010) demonstrated the ability to balance student participation when awarding continuous credit (daily credit) for participation, whereas the current study aimed to produce similar results using random credit. Use of random credit could decrease the amount of instructor grading time, while maintaining balanced participation. In addition, the study examined characteristics of students participating at the various levels via a self-report survey, as well as examining the relationship between critical thinking, exam grades, and student participation levels.

Effect of Credit on Patterns of Student Participation

The current study is an extension of the Krohn et al. (2010) study in which continuous credit was provided for student participation. This study extends the research by comparing the results of randomly selecting credit days within credit units with results previously found under a continuous credit schedule. Krohn et al. showed that continuous credit decreased the percent of non-participants, increased the percent of credit-level participants, maintained the percent of frequent participants, and decreased the percent of dominant participants. Random credit was shown to produce equivalent results for non-participants and credit-level participants as continuous credit did. Neither continuous nor random credit significantly influenced the percent of frequent participants. However, continuous credit was shown to be effective in decreasing the

percent of dominant participants, whereas random credit could not due to a lack of dominant participants in any phase of the study.

While there were many similarities between the effects of continuous and random credit contingencies, two key differences contributed to differences in the effects of the two credit arrangements. First, the baseline percentages of non-participants (0 comments per day) were much higher, ranging from 40 to 58, in the Krohn et al. study (continuous credit) than the current study, which ranged from approximately 16 to 32%. Second, the baseline percent of credit-level participants (1 to 2 comments per day) was much lower (23 to 33%) under continuous credit than random credit (51 to 61%). Therefore, the study using continuous credit had baseline levels that allowed greater improvement than the study using random credit.

After baseline, the percent of non-participants was basically the same between the two studies (approximately 40 to 60% in non-credit units and 20 to 30% in credit units). Likewise, the percent of credit-level participants was similar in both studies (approximately 20 to 30% in non-credit units and 40 to 60% in credit units). Thus, even though there were some pre-existing differences between the two studies, similar changes in the percent of non-participants and credit-level participants were observed in both studies.

The present study was also compared to the Aspiranti (2010) study in which students were given credit at the end of course. When Aspiranti informed the students as to the units in which credit would be provided, analogous percentages of non-participants and credit-level participants were observed as in the current study. The remaining two course sections of Aspiranti's study produced fairly consistent percentages of non-participants across all units. These two sections were not told which units would be selected for credit until the course

concluded. Therefore, it would not be reasonable to expect these two units to have predictably differing amounts of non-participants and credit-level participants across units.

Neither Krohn et al. (2010) nor the current study was effective in greatly changing the level of frequent participants (3 to 4 comments per day) across credit and non-credit phases. The Krohn et al. study had slightly greater differences in mean levels of frequent participants between credit and non-credit phases than did the present study. Also, the Krohn et al. study exhibited numerous overlapping data points in comparisons of frequent participation in credit and non-credit phases. Differences between the mean level of frequent participants across credit and non-credit phases were slight in the present study, except for Section B. Similar to the Krohn et al. study, random credit produced several overlapping data points across credit and non-credit units for frequent participants. Thus, neither study was effective in controlling frequent participants level of involvement in discussion. One possible explanation for this finding is that frequent participants are more likely to be involved in discussion due to personality factors (intrinsic and/or social) than those participating at other levels.

The largest discrepancy between the two studies was the effect of participation on dominant participants (5 or more comments per day). In baseline, Krohn et al. (2010) reported many more dominant participants (mean range approximately 7 to 15%) than the current study which only had dominant participants in one section (mean less than 1%). Therefore, while Krohn et al. had dominant participants in the study who could be influenced by the credit contingency, the present study had virtually none. Krohn et al. demonstrated that continuous credit reduced dominant participants to nearly zero in two of the three course sections. In the third section in which students did not self-record their comments, the percentage of dominant

participants was reduced to approximately 5 to 8% in credit units. However, when dominant participants received credit and self-recorded their comments, dominant participation essentially disappeared. Thus, it may be that self-recording of comments may sensitize dominant participants to the effects of their dominant participation on other students' opportunity to participate. Regardless, the combination of self-recording and a credit contingency nearly eliminated the number of dominant participants in both studies. Also, the lack of dominant participants in the current study may be due to the fact that students under-reported their participation once they achieved credit-level.

Effect of Credit on Low- Medium- and High Participants

Low-participants (mean 0.5 comments or less in baseline) tended to remain low-participants throughout of the course. Approximately 39% of low-participants in baseline remained low-participants across at least four of the five course units and 61% moved to a higher participation category. There were no significant differences between the number of low-participants in credit and non-credit units, which suggests that random credit does not have an effect on initially reticent students.

Medium-participants (0.51 to 1.99 comments per day in baseline) and high-participants (2 or more comments per day in baseline) were similarly tracked. In non-credit units 62% of medium-participants moved to the low-participant category and 11% rose to the high-participant category. During credit units, 21% became low-participants and 32% became high-participants. Of the high-participants, 25% moved into either low- or medium-participant groups during credit units and 67% fell into low- or medium-participant groups during non-credit units. Thus, the

credit contingency appeared to have a greater impact on initially medium- and high-participants than on initially low-participants.

Reliability and Student Participation

Students' self-recorded participation every discussion day, and their participation was recorded by observers on the third discussion day of each unit. Inter-observer agreement between students and observers ranged from 62 to 92%. A pattern of heightened student-observer agreement was observed in credit units. Also, students were significantly more likely to under-report participation in non-credit units than in credit units and students were significantly more likely to over-report participation in credit units than non-credit units. Students' under-reporting of participation has been reported by Aspiranti (2010), Krohn (2010), and Krohn et al. (2010). However, other researchers (Burchfield & Sappington, 1999; Gopinath, 1999) concluded that students over-report their participation in assessments conducted at the end of each class period. Although neither over- nor under-reporting is desirable, over-reporting to gain undue credit is more problematic. Therefore, an instructor wishing to base participation grades on student self-reports should have the students provide self-reports throughout the course rather than at the end of the course. Under-reporting was not likely to have adversely affected grades inasmuch as students only under-reported their participation level once they had participated at the credit-level.

Non-effect of Differential Instructor Behavior

In order to control for the difference of instructor in each section, observers also recorded instructor questions (factual, comprehension, total) and feedback (positive, negative). Across sections, no instructor consistently posed more or fewer comprehension, factual, or total

questions. Although no significant difference was found, the instructor of Section A tended to pose fewer total questions and comprehension questions than the other instructors. Sections A and C did not pose significantly more comprehension or factual questions in credit than non-credit units. However, Section B posed significantly more comprehension and factual questions in non-credit than credit units. A Chi-square analysis revealed no significant differences between instructors across units in their pattern of questions. The slight difference between instructor questions is more likely a result of the group of students and their comments and understanding than a difference in instructor style. A LSD post-hoc analysis showed the instructor of Section B to provide significantly more positive feedback than instructors in Sections A and C, but no interaction effects were found for the amount of negative feedback provided by instructors. The slight differences observed between instructor questions and feedback could be attributable to the instructors posing more questions to explore higher-order concepts due to a particular class's greater grasp of the material. As anyone who has taught multiple courses would likely attest, each group of students develops its own personality, style of communication, and cognitive mastery.

One could claim that credit effects could have been jeopardized if instructors had posed significantly different amounts of questions across sections. However, one could also assert that the number of questions posed does not heighten or limit the number of opportunities students have to participate. For example, one well-designed comprehension question could produce a multitude of comments from a variety of students, whereas a simplistic question may evoke only a few comments.

Hodge and Nelson (1991) documented the impact of positive and negative feedback on student participation rates, showing that students given positive feedback were more likely to participate and those given negative feedback were less likely to participate. In the current study, one section (B) provided significantly more positive feedback than the other two sections. All sections were equivalent in regard to the amount of negative feedback provided. Overall, each section received a high amount of positive feedback and minimal negative feedback. The greater amount of positive feedback provided in Section B compared to Sections A and C is unlikely to have affected the treatment condition and is more likely a result of the personality of the class as noted above. No significant differences between credit and non-credit units were found for the number of negative and positive feedback statements provided by instructors.

Effect of Participation Credit on Quality of Comments

One concern of providing systematic credit for participation in which student's have a perceived "quota" is that students' might provide comments of little value merely to receive credit. One way of addressing this concern is to have students and observers rate each student comment as timely or repetitious. In the current study, a comment was considered timely if it was voluntary, on-topic, and non-redundant. Repetitious comments were voluntary and already posed by another student. To help students judge whether their comment was timely or repetitious, the instructors regularly reviewed the definitions with the students and provided feedback for each comment that indicated whether the comment was repetitious or timely (e.g., "Samuel just asked that question").

As with previous research (Aspiranti, 2010; Krohn, 2010) on the quality of student comments, when students and observers recorded redundant comments, the inter-observer

agreement between observers and students and between observers was poor. However, inter-rater agreement between observers was high when they recorded timely comments. Inter-rater mean percent of agreement between observers and students ranged from 59 to 92% and the mean percent of agreement between observers ranged from 73 to 100% for timely comments. Mean percent of inter-rater agreement for repetitious comments ranged from 0 to 100% between students and observers and between observers.

The low inter-rater agreement for repetitious comments was likely due to the extremely low number of repetitious comments. For example, the most repetitious comments recorded by either observer or student in any unit was seven. No consistent trends were observed across credit contingencies for either timely or repetitious comments. Due to low reliability and low frequency of repetitious comments, no conclusions can be made regarding the effect of credit contingencies on poor-quality comments. However, the fact that there were few repetitious comments recorded suggests that crediting participation does not decrease the quality of discussion.

Further work is needed in assessing quality of comments. While it may be intuitive to think that fewer qualitative categories would produce higher reliability than multiple categories, that pattern may not prove to be the case. Including the current study, three studies (Aspiranti, 2010; Krohn, 2010) have produced weak reliability in their dichotomous (timely or repetitious) assessment of the quality of participation. The dichotomous measure may oversimplify the options presented to students, resulting in less cognitive vigilance as to how to rate comments. Also, students may regard the distinction as trivial or unimportant due to the oversimplification of a complex construct. Creating three to five well-defined categories in which the quality of a

comment can be rated may increase the perceived importance of accurate ratings. Also, providing well-defined categories may help students understand the importance of participating in discussion and the difference between high and poor quality comments.

Participation Survey as a Predictor of Participation

One of the secondary goals of this study was to predict class participation levels using a participation survey adopted from Aspriranti (2010) and Krohn (2010) and determine characteristics of students participating at various levels (low, medium, high). Such information may be useful in planning credit contingencies and classroom environments to balance discussion. Consistent with previous research (Krohn), I found the Participation Survey to be a strong predictor of low and high participation. Average percent of participants correctly classified as low or high was 69% across units and sections. A three-factor model (Personal Benefits, Expectation, History/Confidence) developed by Krohn (2010) was used in the current study. The three-factor model produced an R^2 greater than that of the total survey and on average correctly classified students 75% of the time into low and high participations groups compared to the 69% for the total survey. History/Confidence was the only significant factor contributing to the predictability of the three-factor model and produced nearly equivalent results as those by the three-factor model. Overall, the History/Confidence factor correctly classified students 72% of the time into low- and high- participation groups.

Krohn (2010) also found the History/Confidence factor to predict low- and high-participation levels as well as the total survey and the three-factor model. Thus, it would likely be more efficient for instructors to use the eight-item History/Confidence factor than the three-factor model or the considerably longer 50-item total survey. However, use of the other items

may provide further insight into students' motivation to participate or not. The History/Confidence factor was not a better predictor in credit than non-credit units, which is inconsistent with the results of Krohn who found that it was a better predictor of low participants in non-credit units and of high participants in credit units.

Future studies should explore the differences between those students who increased participation under credit conditions and those who remained reticent. Ascertaining this difference may lead to differential interventions to increase participation among the most reticent students. Even though one's history of participation in courses cannot change, it may be beneficial to focus on making the student comfortable with the classroom and the teacher at the onset of the course. For example, on the first day of class the instructor may have the students form small groups in which each student must learn about the other group members and then have a student present the information learned about another student to the class as a whole. Additionally, in small classes each student could be required to meet briefly with the instructor during office hours in order to build rapport between the two.

Critical Thinking and Exam Grades as Predictors of Participation

Initially, students were classified into low, medium, and high critical thinking groups. Students in the low critical thinking group were significantly more likely to be low participants than high participants. There was no significant difference between the number of students in the high critical thinking group who were low or high participants. Thus, students with low critical thinking scores were less likely to participate whereas students with high critical thinking scores were just as likely to be low or high participants.

In addition, students were placed in either low (bottom quartile) or high (top quartile) critical thinking groups and exam performance groups. Low and high critical thinking groups and low and high exam groups in combination significantly predicted classification of students into low and high participation groups. Overall, students were correctly classified into low and high participation groups 70% of the time based on their combined critical thinking and exam scores, with about 8% more students being correctly placed into the high participation group than low participation group. Critical thinking groups accurately predicted participation groups 67.4% of the time, whereas exam scores correctly predicted participation groups 69.4% of the time. Both critical thinking and exam performance were better predictors of high participants than low participants. These results are consistent with research by McCleary et al. (2011) who found high participants tended to have high critical thinking scores and performed nearly 8 percentage points better on course exams than low participants.

These results may indicate that students disinclined to participate do so for different reasons. For example, students in the low critical thinking group may choose not to participate in class due to cognitive factors (e.g., inability to make sense of the information presented, a perception of having nothing to contribute), whereas students in the high critical thinking group may choose not to participate due to personality/emotional factors (e.g., anxiety, introversion). Again, further research should be done in this area to differentiate the reasons why students choose not to participate. Identifying reasons for students' lack of participation should better inform instructors how best to encourage participation from various students and consequently balance participation across students.

Limitations and Future Directions

One major limitation of the study is its primary focus on quantitative levels of participations with little regard for qualitative measurement of participation. In all four studies (Aspiranti, 2010; Krohn, 2010; Krohn et al., 2010; and the current study) using a similar model to increase participation, efforts to distinguish comments as timely or repetitious were minimally successful. Rare instances of repetitious comments contributed to poor inter-observer agreement. Nonetheless, there is some evidence (Foster, McCleary, & Williams, 2010) that quality and quantity of comments are complimentary. That is, students making the most comments were rated as having higher-quality comments than students contributing fewer comments.

One might doubt whether requiring 54 students to contribute two comments in a 50-minute class is even possible. The overall mean percent of students commenting in class each day was 74%. The present study attempted to alleviate some of this concern by first choosing a student with a raised hand that had not yet contributed or that seldom contributed to class discussion. Reticent students may have rarely if ever attempted to contribute but they may also not have been recognized by the instructor. In addition, it was common in all three sections for an instructor question to not receive a response for several seconds. Thus, it is likely safe to assume that most students had an opportunity to contribute at least twice each class period.

Another limitation of the study was the difference of content difficulty across units. The baseline unit covered material likely to be most familiar to students: physical development (safety, diet, exercise, drug abuse, sexuality). The second unit, cognitive development (intelligence testing, creativity, critical thinking), is most likely the hardest for students, who are

typically not highly exposed to this information before taking the course. Other studies conducting research in the same course have also found students to contribute more in the first unit than the second unit (Aspiranti, 2010; Foster et al., 2010; Krohn et al., 2010). In fact, units in which credit was provided rarely produced more participation than did the baseline unit.

An over-justification or contrast effect may explain the pattern of participation in the current study. The credit units seldom produced results superior to baseline, whereas subsequent non-credit units often precluded results inferior to baseline. This credit for participation may reduce the desire to participate when no credit is available. Examination of Unit 2 showed participation patterns to be roughly equivalent to baseline participation levels in Section A and C when credit was later awarded in those sections. In Section B, Unit 2 was a non-credit unit and participation levels were less balanced than even baseline levels. This contrast indicates the difficulty of Unit 2 compared to Unit 1 and provides evidence that the credit contingency was effective even though the participation levels in Sections A and C appear similar to those in baseline.

Currently, the available research on critical thinking and participation does not lend itself to cause-effect relationships. That is, one cannot say that participating more in class will increase one's cognitive ability or vice versa. A post-test measure of critical thinking in conjunction with participation data would assist in assessing the causal contribution of critical thinking to participants and vice versa (i.e., whether cognitive factors increase participation or participation increases cognitive factors). For example, such information would allow the researcher to examine whether heightened participation results in improved critical thinking scores. In addition, a pre- and post-test measure of personality/emotionality should also be

administered to determine its role in participation.

Rather than focusing almost exclusively on the quantity of participation, future studies should more precisely assess the quality of participation. Qualitative levels of participation will need to be clearly identified for both observers and students. Educating students on what instructors deem valuable to the class will help them evaluate their comments more accurately and hopefully prompt them to increase the amount of critical inquiry they invest in their responses and class preparation. Also, it may even help some reticent students realize that their private thoughts are on track and would be beneficial for the entire class to hear.

In addition, reticent students may be more likely to participate if participation credit is more strongly weighted at the beginning of the course. If this hypothesis proves to be valid, the next step would be to gradually thin the amount of credit given for participation as the semester continues. In this manner, students would be primed to participate throughout the course, as they become more comfortable talking in class and learn that no embarrassing consequences occur as a result.

Even though a participation survey was given to students to discern factors contributing to a student's likelihood of participating in class, the survey results were not utilized in determining credit contingencies or additional ways to promote participation of reticent students. Future studies should explore reasons why students choose not to participate and effective strategies for increasing their participation. Small discussion groups early in the semester that make group members a vital part of each group's success may help students to build bonds with other class members and strengthen their self-esteem. However, these groups will have to be carefully designed and monitored to prevent students from finding ways to remain non-

participants and/or become dominated by their classmates, making it even more unlikely that they will participate in class discussion.

References

Aspiranti, K. B. (2010). *Random credit for participation in college classes*. Unpublished doctoral dissertation, University of Tennessee, Knoxville.

Aitken, J. E., & Neer, M. R. (1993). College student question-asking: The relationship of classroom communication apprehension and motivation. *The Southern Communication Journal, 59*(1), 73-79.

Auster, C. J., & MacRone, M. (1994). The classroom as a negotiated social setting: An empirical study of the effects of faculty members' behavior on students' participation. *Teaching Sociology, 22,* 289-300.

Barnes, C. P. (1983). Questioning in college classrooms. In C. L. Ellner & C. P. Barnes (Eds.), *Studies of college teaching* (pp. 61-82). Lexington, MA: D. C. Heath.

Bean, J. C., & Peterson, D. (1998). Grading classroom participation. *New Directions for Teaching and Learning, 74,* 33-40. doi: 10.1002/tl.7403

Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives, handbook I: Cognitive domain*. New York: David McKay.

Boniecki, K. A., & Moore, S. (2003). Breaking the silence: Using a token economy to reinforce classroom participation. *Teaching of Psychology, 30*, 224-227. doi: 10.1207/s15328023t0p3003_05

Burchfield, C. M., & Sappington, J. (1999). Participation in classroom discussion. *Teaching of Psychology, 26*, 290-291. doi: 10.1207/s15328023t0p3003_05

Christensen, L. J., Curley, K. E., Marquez, E. M., & Menzel, K. E. (1995, November). *Classroom situations which lead to student participation.* Paper presented at the Annual Meeting of Speech Communication Association, San Antonio, TX.

Connor-Greene, P. A. (2005). Fostering meaningful classroom discussion: Student-generated questions, quotations, and talking points. *Teaching of Psychology, 32,* 173-174.

Dallimore, E. J., Hertenstein, J. H., & Platt, M. B. (2004). Classroom participation and discussion effectiveness: Student-generated strategies. *Communication Education, 53*, 103-115. doi: 10.1080/0363452032000135805

Davis, B. G. (1993). *Tools for teaching.* San Francisco: Jossey-Bass.

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research, 59*, 395-430.

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

Fassinger, P. A. (1995b). Professors' and students' perceptions of why students participate in class. *Teaching Sociology, 24,* 25-33. doi:10.2307/1318895

Ferguson, G. A., & Takane, Y. (1989). *Statistical analysis in psychology and education* (6th ed.). New York: McGraw-Hill Publishing.

Ferguson, N. B. L. (1986). Encouraging responsibility, active participation, and critical thinking in general psychology students. *Teaching of Psychology, 13*, 217-218. doi: 10.1207/s15328023top1304_13

Foster, L. N., Krohn, K. R., McCleary, D. F., Miller, K. B., Nalls, M. L., Quillivan, C. C., et al. (2009). Increasing low-verbal students participation in class discussion. *Journal of Behavioral Education, 18,* 173-188. doi: 10.1007/s10864-009-9083-8

Foster, L. N., McCleary, D. F., & Williams, R. L. (2010, May). Reliability and predictive validity of a new qualitative measure of participation in class discussion. In R. L.

Williams (Chair), *Improvement of the quantity and quality of undergraduates' participation in class discussion.* Symposium conducted at the Annual Convention of Association for Behavior Analysis International, San Antonio.

Garside, C. (1996). Look who's talking: A comparison of lecture and group discussion teaching strategies in developing critical thinking skills. *Communication Education, 45,* 212-227. doi: 10.1080/03634529609379050

Garson, G. D. (2008). *Scales and Standard Measures* from Statnotes: Topics in Multivariate Analysis. Retrieved December 28, 2010 from http://faculty.chass.ncsu.edu/garson/pa765/statnote.htm.

Gilson, C. (1994). Of dinosaurs and sacred cows: The grading of classroom participation. *Journal of Management Education, 18,* 227-236.

Gladding, S. T. (2008). *Groups: A counseling specialty* (5th ed.). New Jersey: Merrill Prentice Hall.

Gopinath, C. (1999). Alternatives to instructor assessment of class participation. *Journal of Education for Business, 75,* 10-14. doi: 10.1080/08832329909598983

Hodge, G. K., & Nelson, N. H. (1991). Demonstrating differential reinforcement by shaping classroom participation. *Teaching of Psychology, 18*, 239-241. doi: 10.1207/s15328023top1804_13

Howard, J. R., & Henney, A. L. (1998). Student participation and instructor gender in the mixed-age college classroom. *Journal of Higher Education, 69,* 384-405.

Howard, J. R., James, G. H., III, & Taylor, D. R. (2002). The consolidation of responsibility in the mixed-age college classroom. *Teaching Sociology, 30,* 214-234.

Jacobs, L. C., & Chase, C. I. (1992). *Developing and using tests effectively: A guide for faculty.* San Francisco: Jossey-Bass.

Janzow, F., & Eison, J. (1990). Grades: Their influence on students and faculty. In M. D. Svinicki (Ed.), *The changing face of college teaching. New directions for teaching and learning no. 42.* San Francisco: Jossey Bass.

Jones, R. C. (2008). The "why" of class participation: A question worth asking. *College Teaching, 56*, 59-63. doi: 10.3200/ctch.56.1.59-64

Krohn, K. R. (2010). *The effect of self-recording and contingent credit on the quantity and relevance of student participation in class discussion in large college classes.* Unpublished doctoral dissertation, University of Tennessee, Knoxville.

Krohn, K. R., Aspiranti, K. B., Foster, L. N., McCleary, D. F., Taylor, C. M., Nalls, M. L., Quillivan, C. C., et al. (2010). Effects of self-recording and contingent credit on balancing participation across students. *Journal of Behavioral Education, 19*, 134-155. doi: 10.1007/s10864-010-9105-6

Krohn, K. R., Foster, L. N., McCleary, D. F., Miller, K. B., Nalls, M. L., Quillivan, C. C., et al. (in press). *Reliability of students' self-recorded participation in class discussion*. *Teaching of Psychology*.

Love, K. B. (1981). Comparison of peer assessment methods: Reliability, validity, friendship bias, and user reaction. *Journal of Applied Psychology, 66,* 451-457. doi:10.1037/0021-9010.66.4.451

Mainkar, A. V. (2008). A student-empowered system for measuring and weighing participation in class discussion. *Journal of Management Education, 32,* 23-37. doi: 10.1177/1052562906286912

McCleary, D. F., Foster, L. N., & Williams, R. L. (2011). Relatinship of class participation to critical thinking and test performance. Manuscript submitted for publication.

Melvin, K. B. (1988). Rating class participation: The prof/peer method. *Teaching of Psychology, 15,* 137-139. doi: 10.1207/s15328023top1503_7

Montello, D. R. (1988). Classroom seating location and its effect on course achievement, participation, and attitudes. *Journal of Environmental Psychology, 8*(2), 149-157. doi: 10.1016/S0272-4944(88)80005-7

Renne, T., Kass, H., & Nay, M. A. (1973). The effect of verbalizers on the achievement of non-verbalizers in an enquiring classroom. *Journal of Research in Science Teaching, 10,* 113-124. doi: 10.1002/tea.3660100203

Simpson, E. S., & Courtney, M. (2008). Implementation and evaluation of critical thinking strategies to enhance critical thinking skills in Middle Eastern nurses. *International Journal of Nursing Practice, 14,* 449-454. doi: 10.1111/j.1440-172x.2008.00719.x

Smith, D. G. (1977). College classroom interactions and critical thinking. *Journal of Educational Psychology, 69,* 180-190. doi:10.1037/0022-0663.69.2.180

Sommer, R., & Sommer, B. A. (2007). Credit for comments, comments for credit. *Teaching of Psychology, 34*, 104-106. doi: 10.1207/s15328023top2802_06

Tsui, L. (1998). *A review of the research on critical thinking.* Paper presented at the annual meeting of the Association for the Study of Higher Education, Miami, FL.

Wallace, M. A., & Williams, R. L. (2003). Multiple-choice exams: Explanations for student choices. *Teaching of Psychology, 30,* 136-139.

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

Weaver, R. R., & Qi, J. (2005). Classroom organization and participation: College students' perceptions. *The Journal of Higher Education, 76,* 570-601. doi: 10.1353/jhe.2005.0038

Williams, R. L., Bain, S. K., Stockdale, S. L., Isaacs, R., & Williams, A. (2003). Role of critical thinking in judging the accuracy and sources of claims in human development. *Inquiry: Critical Thinking across the Disciplines, 22,* 65-72.

Williams, R. L., Oliver, R., Allin, J. L., Winn, B., & Booher, C. S. (2003a). Psychological critical thinking as a course predictor and outcome variable. *Teaching of Psychology, 30,* 220-223. doi: 10.1207/S15328023TOP3003_04

Williams, R. L., Oliver, R., Allin, J. L., Winn, B., & Booher, C. S. (2003b). Knowledge and critical thinking as course predictors and outcomes. *Inquiry: Critical Thinking across the Disciplines, 22,* 57-63.

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. *Journal of General Education, 53*, 37-58. doi: 10.1353/jge.2004.0022

Williams, R. L., & Stockdale, S. L. (2003). High-performing students with low critical thinking skills. *Journal of General Education, 52,* 200-226. doi: 10.1353/jge.2004.0007

Williams, R. L., & Worth, S. L. (2002). Thinking skills and work habits: Contributors to course performance. *The Journal of General Education, 51,* 200-227. doi: 10.1353/jge.2003.0007

Appendices

Appendix A: Tables

Table 1

Credit Contingency per Unit

	Unit				
	1	2	3	4	5
Section A (12:20)	Baseline	RC^{2x}	NC	RC^{2x}	NC
Section B (10:10)	Baseline	NC	RC^{2x}	NC	RC^{2x}
Section C (11:15)	Baseline	RC^{2x}	NC	RC^{2x}	NC

Note. NC = non-credit; RC^{2x} = random credit for two days within unit.

Table 2

Mean Daily Participation Level Based on Student Gender and Academic Classification

	Mean Daily Participation		
	Overall	NC	RC^{2x}
Gender (*n*)			
Male (31)	1.25 (.81)	.86 (.88)	1.43 (.90)
Female (130)	1.48 (.96)	1.12 (1.07)	1.71 (1.06)
Academic Classification (*n*)			
Freshman (2)	.53 (.38)	.32 (.25)	.81 (.62)
Sophomore (71)	1.37 (.85)	.95 (.94)[a]	1.61 (.96)
Junior (57)	1.35 (.95)	.99 (1.03)	1.56 (1.08)
Senior (21)	1.60 (1.00)	1.28 (1.12)	1.76 (1.07)
Graduate (8)	2.14 (.99)	1.93 (1.23)[a]	2.39 (1.08)

Note. Two students did not report academic classification. Overall = daily participation mean computed across all 5 units; NC = non-credit; RC^{2x} = random credit for two days within unit. [a] = significant difference ($p < .05$) between sophomore and graduate non-credit mean participation.

Table 3

Mean Percent of Inter-observer Agreement between Student and Observer Records of Class Participation on Inter-rater Check Days

	Units				
Pairs within sections	1	2	3	4	5
Section A	b	c	nc	c	nc
Students and observer 1	80 (29)	86 (35)	71 (18)	80 (26)	67 (15)
Students and observer 2	80 (29)	79 (36)	70 (18)	80 (26)	71 (15)
Observers 1 and 2	94 (26)	88 (33)	96 (17)	96 (23)	95 (15)
Section B	b	nc	c	nc	c
Students and observer 1	89 (37)	84 (28)	92 (40)	84 (25)	90 (27)
Students and observer 2	89 (37)	87 (28)	92 (41)	79 (25)	91 (27)
Observers 1 and 2	90 (37)	94 (28)	89 (42)	94 (24)	99 (27)
Section C	b	c	nc	c	nc
Students and observer 1	87 (37)	83 (33)	67 (16)	86 (34)	62 (13)
Students and observer 2	85 (37)	83 (33)	64 (17)	83 (34)	65 (13)
Observers 1 and 2	94 (37)	98 (30)	93 (16)	88 (32)	96 (13)
All sections					
Student and observer 1	86 (103)	84 (96)	82 (74)	84 (85)	77 (55)
Student and observer 2	85 (103)	83 (97)	81 (76)	81 (85)	80 (55)
Observers 1 and 2	92 (100)	93 (91)	91 (75)	92 (79)	97 (55)

Table 4

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]		[c]		[nc]		[c]		[nc]	
Students	1.24	(1.23)	1.73	(1.29)	1.00	(1.32)	1.39	(1.35)	.69	(1.17)
Observer 1	1.13	(1.28)	1.60	(1.50)	1.24	(1.77)	1.30	(1.55)	1.03	(1.86)
Observer 2	1.13	(1.29)	1.53	(1.41)	1.20	(1.74)	1.30	(1.53)	1.05	(1.81)
Section B	[b]		[nc]		[c]		[nc]		[c]	
Students	1.67	(1.24)	1.45	(1.49)	1.53	(1.04)	1.04	(1.14)	1.49	(1.17)
Observer 1	1.79	(1.34)	1.76	(1.83)	1.57	(1.12)	1.26	(1.67)	1.59	(1.46)
Observer 2	1.69	(1.31)	1.64	(1.69)	1.57	(1.04)	1.23	(1.66)	1.56	(1.41)
Section C	[b]		[c]		[nc]		[c]		[nc]	
Students	1.47	(1.12)	1.47	(1.06)	.76	(1.14)	1.45	(1.23)	.49	(.87)
Observer 1	1.68	(1.32)	1.51	(1.29)	1.00	(1.56)	1.30	(1.32)	.78	(1.32)
Observer 2	1.64	(1.33)	1.47	(1.25)	1.02	(1.49)	1.26	(1.19)	.76	(1.30)
All Sections										
Students	1.23	(1.23)	1.55	(1.28)	1.12	(1.2)	1.29	(1.25)	.90	(1.16)
Observer 1	1.13	(1.28)	1.62	(1.54)	1.29	(1.49)	1.28	(1.48)	1.13	(1.59)
Observer 2	1.13	(1.29)	1.54	(1.45)	1.28	(1.44)	1.26	(1.46)	1.13	(1.55)

Table 5

Mean Percent of Inter-Observer Agreement between Student and Observer Records of Timely Student Participation on Inter-rater Check Days

	Units				
Pairs within sections	1	2	3	4	5
Section A	(*n*)[b]	(*n*)[c]	(*n*)[nc]	(*n*)[c]	(*n*)[nc]
Students and observer 1	73 (28)	92 (33)	70 (18)	77 (25)	77 (15)
Students and observer 2	88 (27)	92 (33)	62 (18)	80 (25)	77 (15)
Observers 1 and 2	81 (26)	100 (32)	89 (18)	86 (23)	90 (15)
Section B	(*n*)[b]	(*n*)[nc]	(*n*)[c]	(*n*)[nc]	(*n*)[c]
Students and observer 1	89 (37)	78 (29)	90 (39)	84 (24)	91 (27)
Students and observer 2	87 (37)	86 (28)	87 (40)	83 (24)	91 (27)
Observers 1 and 2	92 (36)	90 (28)	97 (38)	98 (23)	96 (27)
Section C	(*n*)[b]	(*n*)[c]	(*n*)[nc]	(*n*)[c]	(*n*)[nc]
Students and observer 1	79 (37)	83 (31)	59 (18)	80 (34)	59 (12)
Students and observer 2	87 (37)	91 (31)	60 (17)	84 (34)	65 (11)
Observers 1 and 2	84 (37)	92 (30)	74 (19)	85 (32)	73 (12)
All sections	(*n*)	(*n*)	(*n*)	(*n*)	(*n*)
Student and observer 1	81 (102)	85 (93)	78 (75)	80 (83)	80 (54)
Student and observer 2	87 (101)	90 (92)	74 (75)	82 (83)	81 (53)
Observers 1 and 2	86 (99)	94 (90)	89 (75)	89 (78)	90 (54)

Table 6

Mean Percent of Inter-Observer Agreement between Student and Observer Records of Repetitious Student Commenting on Baseline (b), Credit (c), and Non-Credit (nc) Units

	Units				
	1	2	3	4	5
Section A	(*n*)[b]	(*n*)[c]	(*n*)[nc]	(*n*)[c]	(*n*)[nc]
Students and observer 1	0 (8)	0 (4)	13 (4)	0 (3)	NR
Students and observer 2	29 (7)	0 (4)	10 (5)	0 (3)	NR
Observers 1 and 2	0 (5)	NR	50 (2)	NR	NR
Section B	(*n*)[b]	(*n*)[nc]	(*n*)[c]	(*n*)[nc]	(*n*)[c]
Students and observer 1	17 (6)	50 (2)	33 (3)	NR	0 (1)
Students and observer 2	13 (8)	33 (3)	33 (3)	NR	0 (1)
Observers 1 and 2	71 (7)	50 (2)	100 (2)	NR	NR
Section C	(*n*)[b]	(*n*)[c]	(*n*)[nc]	(*n*)[c]	(*n*)[nc]
Students and observer 1	0 (6)	0 (6)	33 (3)	0 (3)	50 (1)
Students and observer 2	0 (3)	25 (4)	20 (5)	0 (4)	0 (1)
Observers 1 and 2	8 (6)	0 (3)	20 (5)	0 (3)	0 (1)

Note. [b] = baseline unit (no credit or self-recording), [c] = credit unit, [nc] = non-credit unit, NR = none recorded.

Table 7

Percent Agreement between Observer Records of Frequency of Instructor Questions on the Inter-rater Check Days in Each Unit

	Units				
	1	2	3	4	5
Section A	[b]	[c]	[nc]	[c]	[nc]
Factual Questions	100	94	86	86	75
Comprehension Questions	95	94	94	97	92
Total Questions	96	94	92	100	86
Section B	[b]	[nc]	[c]	[nc]	[c]
Factual Questions	78	69	83	82	85
Comprehension Questions	89	100	100	94	95
Total Questions	97	97	96	98	91
Section C	[b]	[c]	[nc]	[c]	[nc]
Factual Questions	67	83	86	90	80
Comprehension Questions	92	95	82	91	83
Total Questions	76	97	89	90	94

Note. [b] = baseline units, [c] = credit units, [nc] = non-credit units.

Table 8

Mean Percent of Inter-Observer Agreement between Observers' Records of Instructor Feedback to Students on Inter-rater Check Days in Each Unit

	Units				
	1	2	3	4	5
Section A	$(n)^{b}$	$(n)^{c}$	$(n)^{nc}$	$(n)^{c}$	$(n)^{nc}$
Positive feedback	90 (27)	84 (34)	90 (18)	88 (24)	91 (15)
Negative feedback	0 (3)	50 (4)	0 (2)	25 (8)	40 (5)
Section B	$(n)^{b}$	$(n)^{nc}$	$(n)^{c}$	$(n)^{nc}$	$(n)^{c}$
Positive feedback	87 (37)	92 (29)	87 (42)	92 (24)	96 (28)
Negative feedback	50 (6)	0 (4)	40 (5)	0 (3)	20 (5)
Section C	$(n)^{b}$	$(n)^{c}$	$(n)^{nc}$	$(n)^{c}$	$(n)^{nc}$
Positive feedback	90 (39)	89 (30)	87 (17)	83 (32)	96 (13)
Negative feedback	13 (8)	13 (8)	0 (4)	14 (7)	0 (4)

Note. ^{b} = baseline unit, ^{c} = credit unit, ^{nc} = non-credit unit.

Table 9

Differences in Proportions of Students at Different Participation Levels between Treatment Conditions Overall and between Pairs of Adjacent Treatment Units in Each Section

	Comparisons		
	Overall	Treatment Pair 1	Treatment Pair 2
Section A	(C vs. NC)	(C vs. NC)	(C vs. NC)
0 level	15/46 < 22/43; $p = .04$	14/47 < 25/40; $p = .001$	15/44 < 27/43; $p = .003$
1-2 level	22/46 > 14/43; $p = .07$	24/47 > 9/40; $p = .003$	20/44 > 11/43; $p = .03$
3-4 level	8/46 > 6/43; $p = .33$	8/47 > 6/40; $p = .40$	8/44 > 4/43; $p = .11$
5+ level	1/46 < 1/43; $p = .48$	1/47 > 0/40; $p = .18$	1/44 < 1/43; $p = .49$
Section B	(C vs. NC)	(NC vs. C)	(NC vs. C)
0 level	11/43 < 17/46; $p = .12$	20/45 > 9/44; $p = .008$	22/44 > 13/44; $p = .02$
1-2 level	22/43 > 20/46; $p = .23$	17/45 < 22/44; $p = .12$	16/44 < 23/44; $p = .07$
3-4 level	10/43 > 9/46; $p = .34$	8/45 < 13/44; $p = .10$	6/44 < 8/44; $p = .28$
5+ level	0/43 = 0/46; NP	0/45 = 0/44; NP	0/44 = 0/44; NP
Section C	(C vs. NC)	(C vs. NC)	(C vs. NC)
0 level	11/46 < 21/40; $p = .003$	12/47 < 23/40; $p = .001$	11/46 < 30/39; $p = .0000007$
1-2 level	29/46 > 15/40; $p = .009$	30/47 > 13/40; $p = .002$	28/46 > 8/39; $p = .00004$
3-4 level	6/46 > 4/40; $p = .33$	5/47 < 5/40; $p = .39$	7/46 > 1/39; $p = .03$
5+ level	0/46 = 0/40; NP	0/47 = 0/40; NP	0/46 = 0/39; NP

Note. C = credit; NC = no credit*;* NP = due to both proportions being zero no proportion was calculated.

Table 10

Number of Low Participants in Unit 1 Who Fell into Low, Medium, and High Categories in Subsequent Units

		Units				
		1	2	3	4	5
Section A	n =	15 [b]	15 [c]	15 [nc]	15 [c]	15 [nc]
Low		15	9	15	12	15
Medium		0	6	0	2	0
High		0	0	0	1	0
Section B	n =	7 [b]	7 [nc]	7 [c]	7 [nc]	7 [c]
Low		7	7	6	7	6
Medium		0	0	1	0	1
High		0	0	0	0	0
Section C	n =	11 [b]	11 [c]	11 [nc]	11 [c]	11 [nc]
Low		11	7	10	7	11
Medium		0	4	1	3	0
High		0	0	0	1	0

Note. [b] = baseline units, [c] = credit units, [nc] = non-credit units.

Table 11

Number of Medium Participants in Unit 1 Who Fell into Low, Medium, and High Categories in Subsequent Units

		Units				
		1	2	3	4	5
Section A	n =	17 [b]	17 [c]	16 [nc]	15 [c]	15 [nc]
Low		0	5	8	2	9
Medium		17	6	6	6	5
High		0	6	2	7	1
Section B	n =	21 [b]	21 [nc]	21 [c]	21 [nc]	20 [c]
Low		0	12	1	16	7
Medium		21	7	13	3	7
High		0	2	7	2	6
Section C	n =	18 [b]	18 [c]	17 [nc]	18 [c]	18 [nc]
Low		0	4	9	4	13
Medium		18	10	5	7	3
High		0	4	3	7	2

Note. [b] = baseline units, [c] = credit units, [nc] = non-credit units.

Table 12

Number of High Participants in Unit 1 Who Fell into Low, Medium, and High Categories in Subsequent Units

		Units				
		1	2	3	4	5
Section A	n =	17 [b]	17 [c]	17 [nc]	17 [c]	17 [nc]
Low		0	0	2	1	7
Medium		0	1	5	5	6
High		17	16	10	11	4
Section B	n =	23 [b]	23 [nc]	23 [c]	23 [nc]	23 [c]
Low		0	4	1	4	2
Medium		0	9	2	10	5
High		23	10	20	9	16
Section C	n =	21 [b]	21 [c]	21 [nc]	21 [c]	21 [nc]
Low		0	0	9	2	14
Medium		0	7	5	4	5
High		21	14	7	15	2

Note. [b] = baseline units, [c] = credit units, [nc] = non-credit units.

Table 13

Number of Timely and Repetitious Comments per Unit Recorded by Students and Observers

	Unit 1		Unit 2		Unit 3		Unit 4		Unit 5	
	T	R	T	R	T	R	T	R	T	R
Section A	[b]		[c]		[nc]		[c]		[nc]	
Student	27, 6		33, 4		17, 4		25, 3		14, 0	
Observer 1	23, 2		32, 0		18, 1		33, 0		15, 0	
Observer 2	25, 3		32, 0		17, 2		23, 0		25, 0	
Section B	[b]		[nc]		[c]		[nc]		[c]	
Student	37, 2		28, 2		38, 2		24, 0		27, 1	
Observer 1	36, 5		27, 1		37, 2		23, 0		27, 0	
Observer 2	36, 7		27, 2		38, 2		23, 0		27, 0	
Section C	[b]		[c]		[nc]		[c]		[nc]	
Student	37, 1		31, 4		13, 2		33, 2		11, 1	
Observer 1	34, 5		29, 2		18, 5		31, 1		12, 1	
Observer 2	35, 2		30, 1		16, 7		31, 2		11, 0	

Note. [b] = baseline units, [c] = credit units, [nc] = non-credit units, [sr] = self-record units, [nsr] = non self-record units.

T – Timely comments.

R = Repetitious Comments.

Table 14

Number of Questions (Factual, Comprehension, Total) Posed by Instructors by Unit within Each Section

	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5
Section A	[b]	[c]	[nc]	[c]	[nc]
Factual	4	16	6	6	9
Comprehension	19	16	16	34	23
Total	23	32	22	40	31
Section B	[b]	[nc]	[c]	[nc]	[c]
Factual	7	16	6	14	9
Comprehension	28	26	19	32	19
Total	35	42	25	46	28
Section C	[b]	[c]	[nc]	[c]	[nc]
Factual	6	10	6	20	10
Comprehension	22	20	22	32	15
Total	28	30	28	52	25

Note. [b] = baseline unit, [c] = credit unit, [nc] = non-credit unit.

Table 15

Means and Standard Deviations for Positive and Negative Instructor Feedback following Student Participation per 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]		[c]		[nc]		[c]		[nc]	
Positive	1.14	(1.28)	1.51	(1.38)	1.36	(2.06)	1.25	(1.51)	1.27	(2.42)
Negative	.02	(.14)	.04	(.20)	.02	(.15)	.08	(.28)	.09	(.36)
Section B	[b]		[nc]		[c]		[nc]		[c]	
Positive	1.63	(1.31)	1.63	(1.68)	1.56	(1.07)	1.34	(1.76)	1.67	(1.48)
Negative	.06	(.24)	.00	(.00)	.04	(.19)	.04	(.20)	.02	(.15)
Section C	[b]		[c]		[nc]		[c]		[nc]	
Positive	1.69	(1.42)	1.60	(1.41)	1.22	(1.75)	1.40	(1.65)	1.24	(2.28)
Negative	.04	(.19)	.11	(.31)	.00	(.00)	.08	(.27)	.00	(.00)
Combined										
Positive	1.49	(1.35)	1.58	(1.48)	1.38	(1.65)	1.33	(1.64)	1.39	(2.09)
Negative	.04	(.20)	.05	(.22)	.02	(.14)	.07	(.25)	.04	(23)

Note. [b] = baseline unit, [c] = credit unit, [nc] = non-credit unit.

Table 16

Logistic Regression Results using the Total Participation Survey Items to Predict Placement in Low and High Participation Groups each Unit

Unit	*B*	SE	Wald statistic	Odds Ratio	95% CI for Odds Ratio	
					Lower	Upper
1[a]	.04	.01	11.08***	1.04	1.00	1.06
2[b]	.05	.01	12.73***	1.05	1.02	1.08
3[c]	.03	.01	6.79**	1.03	1.01	1.06
4[d]	.03	.01	6.03*	1.03	1.01	1.05
5[e]	.03	.01	4.41*	1.04	1.00	1.06

[a]Significance for the full model, X^2 (1) = 12.96, $p < .001$, $R^2 = .14$.
[b]Significance for the full model, X^2 (1) = 15.43, $p < .001$, $R^2 = .14$.
[c]Significance for the full model, X^2 (1) = 7.55, $p < .01$, $R^2 = .07$.
[d]Significance for the full model, X^2 (1) = 6.61, $p < .05$, $R^2 = .07$.
[e]Significance for the full model, X^2 (1) = 4.74, $p < .05$, $R^2 = .04$.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 17

Percent of Cases Correctly Classified into Low and High Groups each Unit by Section with the Total Participation Survey Score for Each Student as the Predictor Variable

	Unit				
	1	2	3	4	5
Section A	[b]	[c]	[nc]	[c]	[nc]
Low Participants	75.0	66.7	97.0	66.7	100
High Participants	71.4	78.9	11.1	62.5	0
Overall	73.3	73.0	78.6	64.7	86.5
Section B	[b]	[nc]	[c]	[nc]	[c]
Low Participants	44.4	82.3	22.2	92.3	68.8
High Participants	90.0	36.4	100.0	33.3	64.3
Overall	75.9	68.6	75.9	81.3	66.7
Section C	[b]	[c]	[nc]	[c]	[nc]
Low Participants	0.0	50.0	100.0	0.0	100.0
High Participants	100.0	73.3	0.0	100.0	0.0
Overall	58.6	62.1	83.8	57.6	97.9
All Sections Combined					
Low Participants	45.9	71.4	93.2	84.5	100
High Participants	82.4	60.0	14.3	34.1	0
Overall	67.0	66.3	67.6	63.6	81.8

Note. [b] = baseline unit, [c] = credit unit, [nc] = non-credit unit.

Table 18

Logistic Regression Results using the Participation Survey Factors as Predictors in each Unit

Unit Factor	*B*	SE	Wald statistic	Odds Ratio	95% CI for Odds Ratio Lower	Upper
Unit 1[a]						
History and Confidence	.16	.07	6.20*	1.18	1.04	1.34
Expectation	-.05	.09	0.26	0.95	0.80	1.14
Personal Benefits	.12	.09	1.97	01.3	0.95	1.33
Unit 2[b]						
History and Confidence	.24	.07	11.63**	1.27	1.11	1.45
Expectation	-.05	.09	0.27	0.95	0.79	1.44
Personal Benefits	.05	.08	0.41	1.05	0.90	1.24
Unit 3[c]						
History and Confidence	.15	.07	5.281*	1.17	1.02	1.33
Expectation	-.05	.09	0.32	0.95	0.81	1.13
Personal Benefits	.07	.08	0.78	1.07	0.92	1.24
Unit 4[d]						
History and Confidence	.16	.06	6.75**	1.17	1.04	1.31
Expectation	-.03	.08	0.15	0.97	0.83	1.13
Personal Benefits	.01	.08	0.03	1.01	0.87	1.18
Unit 5[e]						
History and Confidence	.10	.07	1.71	1.10	0.95	1.27
Expectation	.05	.10	0.24	1.05	0.87	1.27
Personal Benefits	-.01	.09	0.00	1.01	0.85	1.19

[a]Significance for the full model, X^2 (3) = 21.15, $p < .001$, $R^2 = .21$.
[b]Significance for the full model, X^2 (3) = 27.80, $p < .001$, $R^2 = .24$.
[c]Significance for the full model, X^2 (3) = 12.97, $p < .01$, $R^2 = .11$.
[d]Significance for the full model, X^2 (3) = 14.60, $p < .01$, $R^2 = .14$.
[e]Significance for the full model, X^2 (3) = 7.85, $p < .05$, $R^2 = .11$.
* $p < .05$. ** $p < .01$.

Table 19

Percent of Cases Correctly Classified into Low and High Groups for the Combined Three Survey Factors as a Model for Predicting Participation

	Unit				
	1	2	3	4	5
Section A	[b]	[c]	[nc]	[c]	[nc]
Low Participants	75.0	77.8	93.9	66.7	93.8
High Participants	78.6	84.2	44.4	68.8	60.0
Overall	76.7	81.1	83.3	67.6	89.2
Section B	[b]	[nc]	[c]	[nc]	[c]
Low Participants	44.4	87.5	33.3	96.2	62.5
High Participants	95.0	72.7	90.0	66.7	57.1
Overall	79.3	82.9	72.4	90.6	60.0
Section C	[b]	[c]	[nc]	[c]	[nc]
Low Participants	25.0	71.4	100.0	50.0	100.0
High Participants	76.5	80.0	0.0	84.2	0.0
Overall	55.2	75.9	83.8	69.7	97.7
All Sections Combined					
Low Participants	54.1	78.6	94.5	79.3	100.0
High Participants	86.3	75.6	25.7	58.5	0.0
Overall	72.7	77.2	72.2	70.7	81.8

Note. [b] = baseline unit, [c] = credit unit, [nc] = non-credit unit.

Table 20

Logistic Regression Results using the History/Confidence Participation Survey Factor as the Predictor in each Unit

Unit Factor	*B*	SE	Wald statistic	Odds Ratio	95% CI for Odds Ratio Lower	Upper
Unit 1[a]						
History and Confidence	.19	.05	14.57***	1.20	1.09	1.32
Unit 2[b]						
History and Confidence	.23	.05	19.25***	1.25	1.13	1.38
Unit 3[c]						
History and Confidence	.14	.05	10.01**	1.15	1.05	1.26
Unit 4[d]						
History and Confidence	.14	.04	11.81**	1.15	1.06	1.25
Unit 5[e]						
History and Confidence	.13	.05	6.35*	1.14	1.03	1.25

[a]Significance for the full model, X^2 (1) = 18.84, $p < .001$, $R^2 = .19$.
[b]Significance for the full model, X^2 (1) = 27.36, $p < .001$, $R^2 = .24$.
[c]Significance for the full model, X^2 (1) = 12.17, $p < .001$, $R^2 = .11$.
[d]Significance for the full model, X^2 (1) = 14.45, $p < .001$, $R^2 = .14$.
[e]Significance for the full model, X^2 (1) = 7.47, $p < .01$, $R^2 = .07$.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 21

Percent of Cases Correctly Classified into Low and High Groups for the History/Confidence Survey Factor as a Model for Predicting Participation

	Unit				
	1	2	3	4	5
Section A	[b]	[c]	[nc]	[c]	[nc]
Low Participants	75.0	66.7	90.9	72.2	96.9
High Participants	71.4	73.7	22.2	56.3	20.0
Overall	73.3	70.3	76.2	64.7	86.5
Section B	[b]	[nc]	[c]	[nc]	[c]
Low Participants	44.4	87.5	33.3	96.2	75.0
High Participants	90.0	72.7	95.0	33.3	50.0
Overall	75.9	82.9	75.9	84.4	63.3
Section C	[b]	[c]	[nc]	[c]	[nc]
Low Participants	41.7	64.3	100.0	42.9	100.0
High Participants	88.2	80.0	0.0	89.5	00.0
Overall	69.0	72.4	83.8	69.7	97.7
All Sections Combined					
Low Participants	45.9	73.2	90.4	79.3	98.9
High Participants	80.4	75.6	25.7	58.5	00.0
Overall	65.9	74.3	69.4	70.7	80.9

Note. [b] = baseline unit, [c] = credit unit, [nc] = non-credit unit.

Appendix B: Figures

Student Record Card

Name: _____ N Card: (□) Yes□ or No□
Unit: _____ Date: _____ IN HW: (□) Yes□ or No□
AQ HW: (□) Yes□ or No□

Voluntary Comments: Check each comment as Timely (T) or Repetitious (R).
1. T ( ) or R ( ) --

2. T ( ) or R ( ) --

3. T ( ) or R ( ) --

Over: (□) Yes□ or No□

Figure 1. Record card for recording participation and daily credit activities.

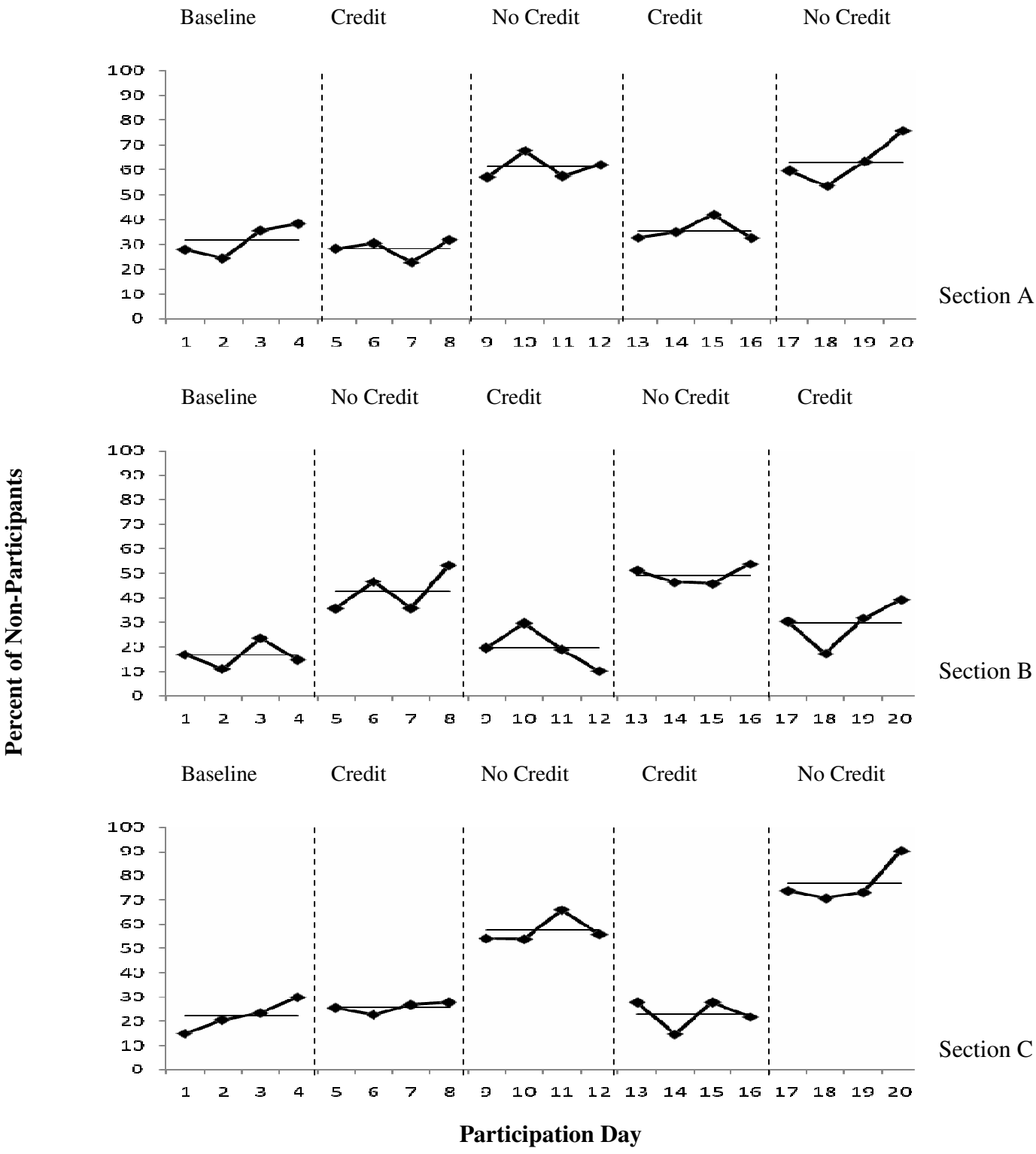


*Figure 2.*Percent of non-participants (0 comment level) each day.

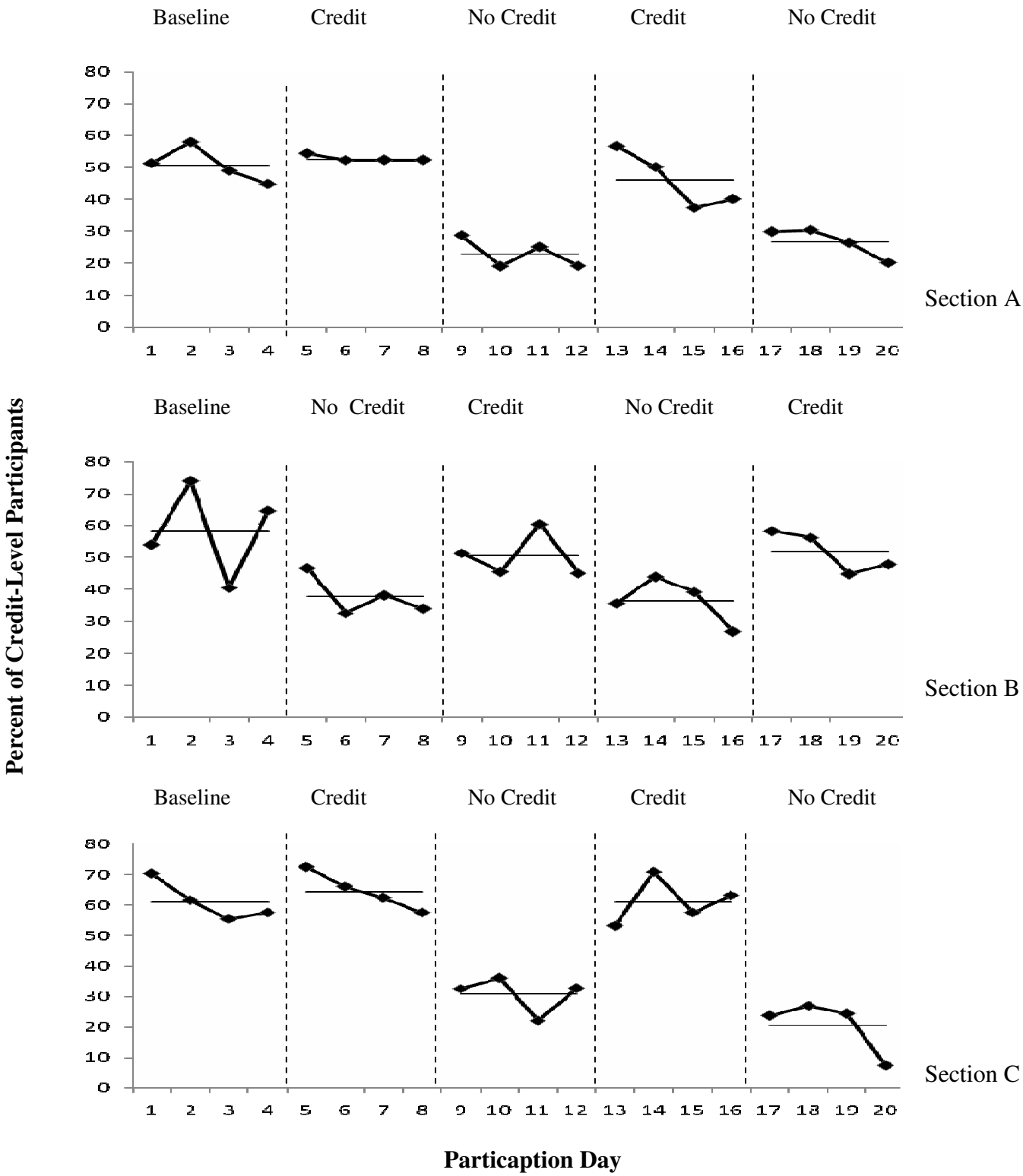


Figure 3. Percent of Credit-Level Participants per day (1-2 comments only) in each unit.

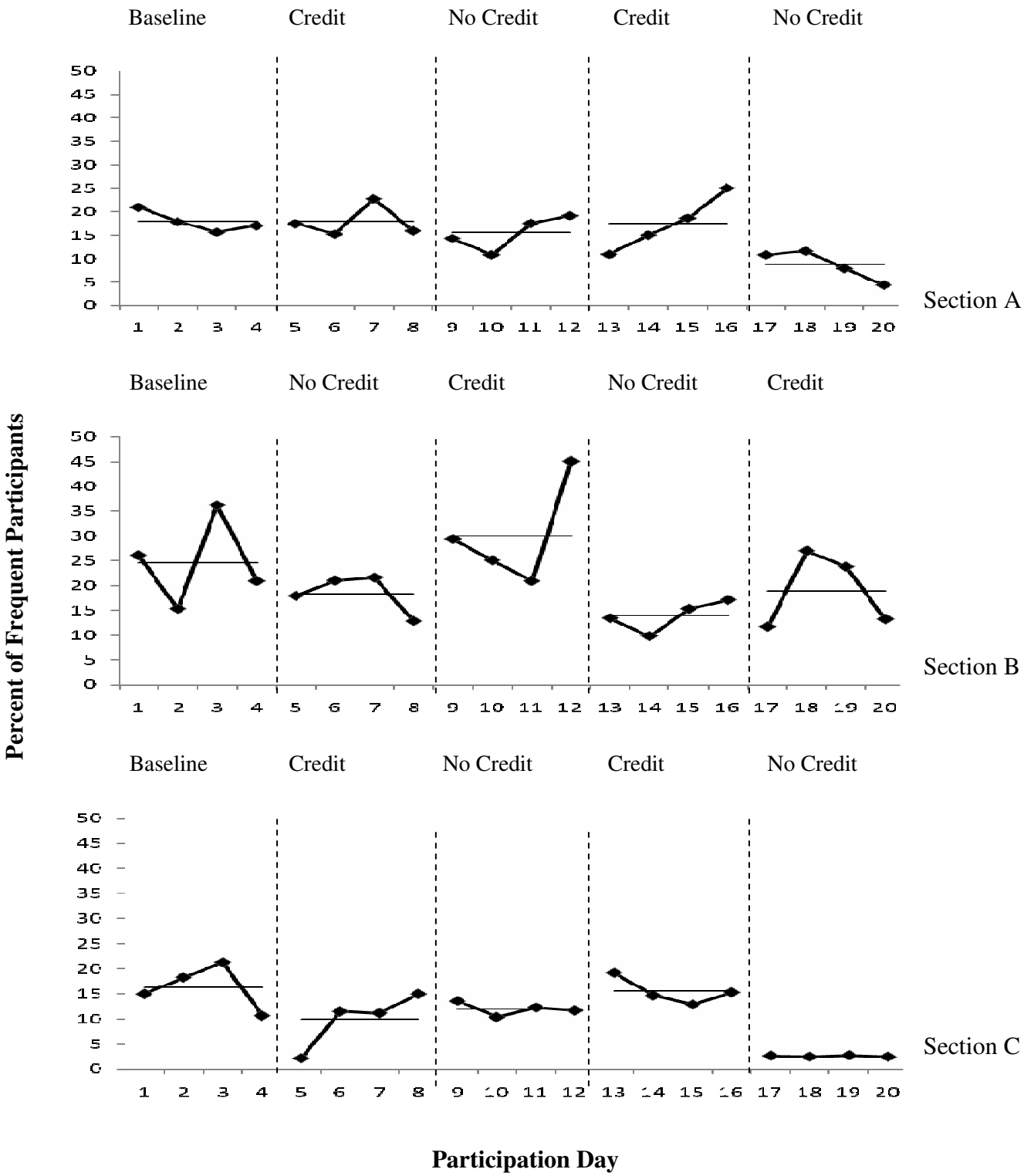


Figure 4. Percent of frequent participants (3-4 comment level) each day.

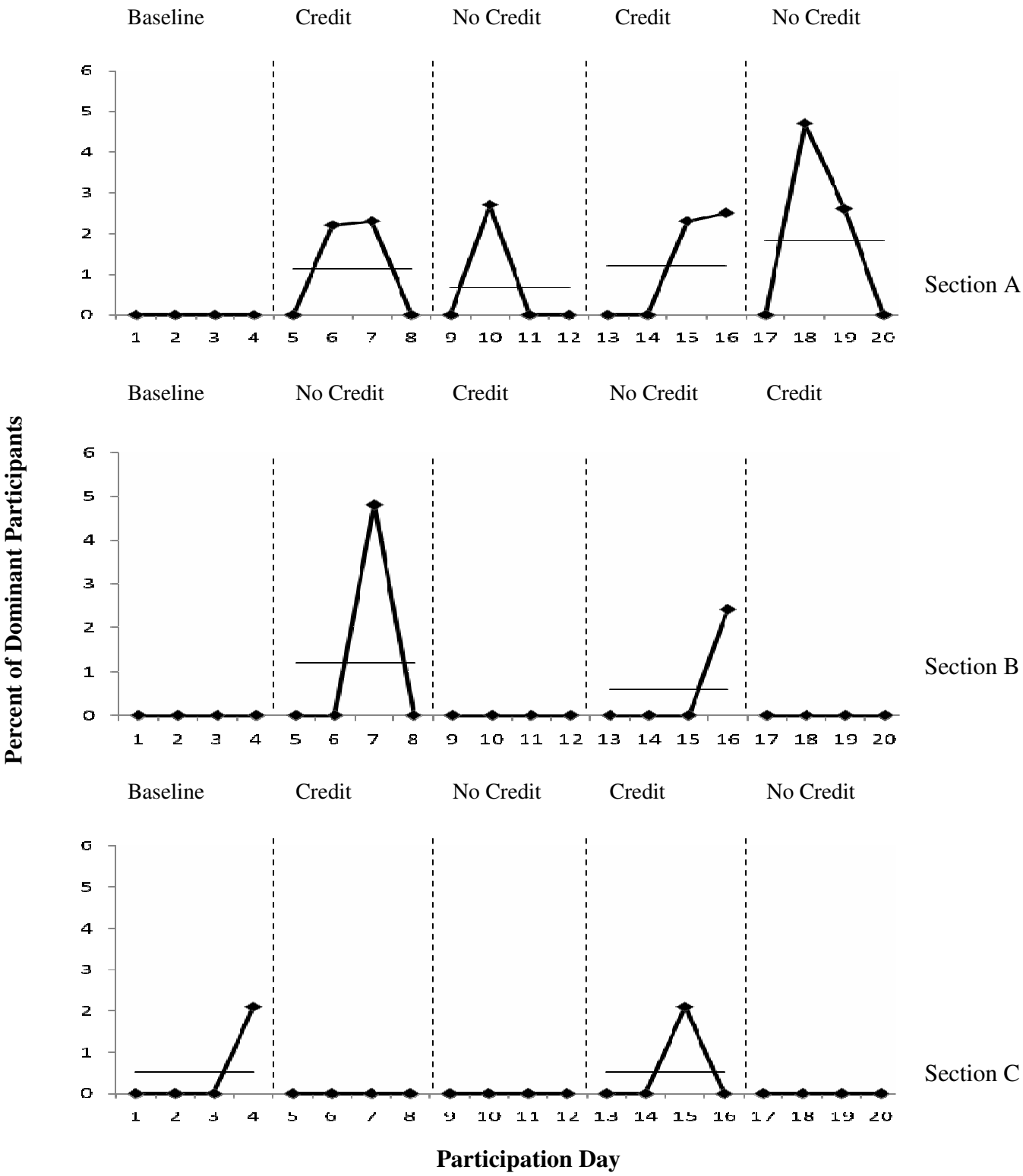


Figure 5. Percent of dominant participants (5+ comment level) each day.

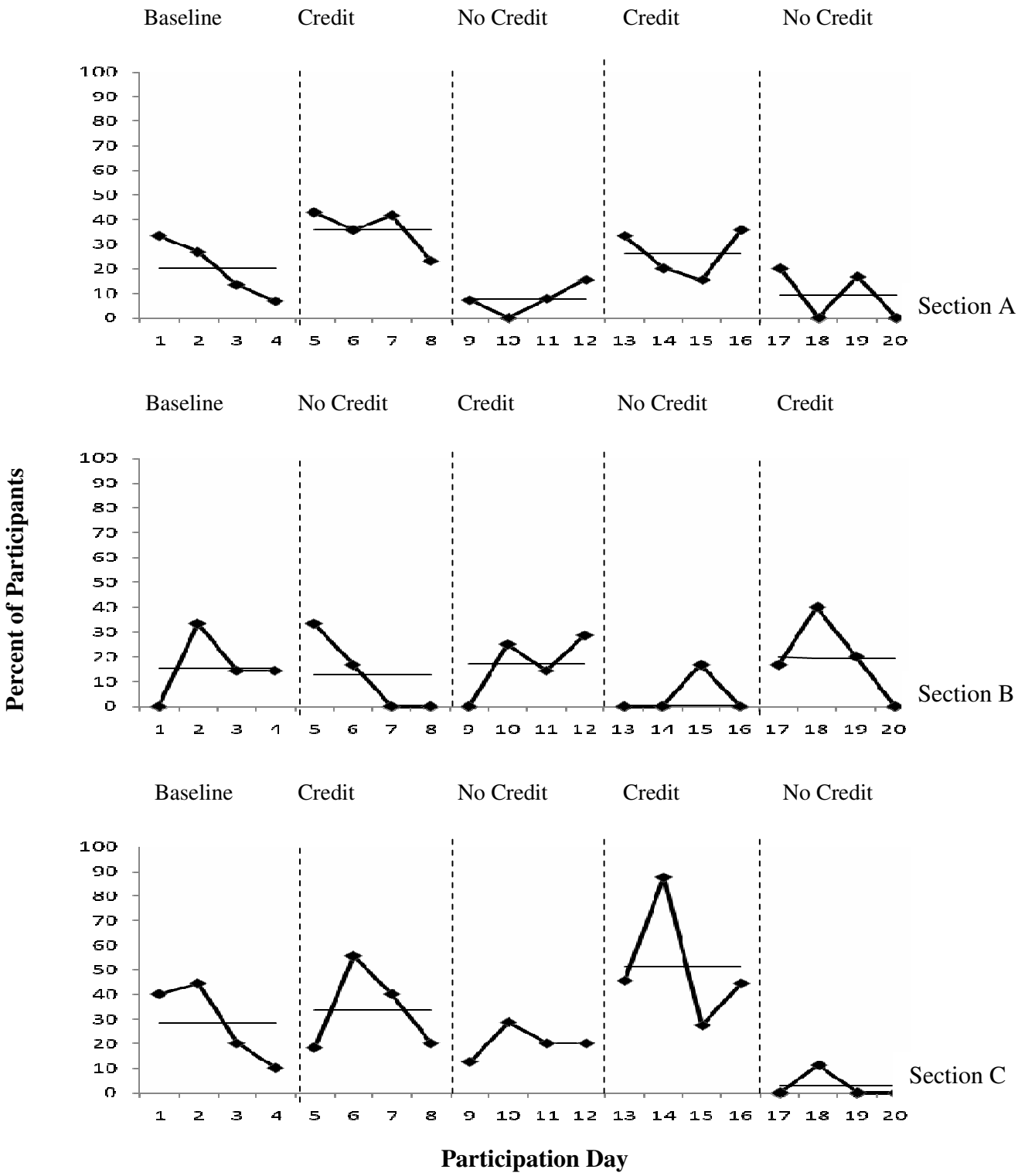


Figure 6. Percent of low-responding participants in Unit 1 who participated in subsequent units.

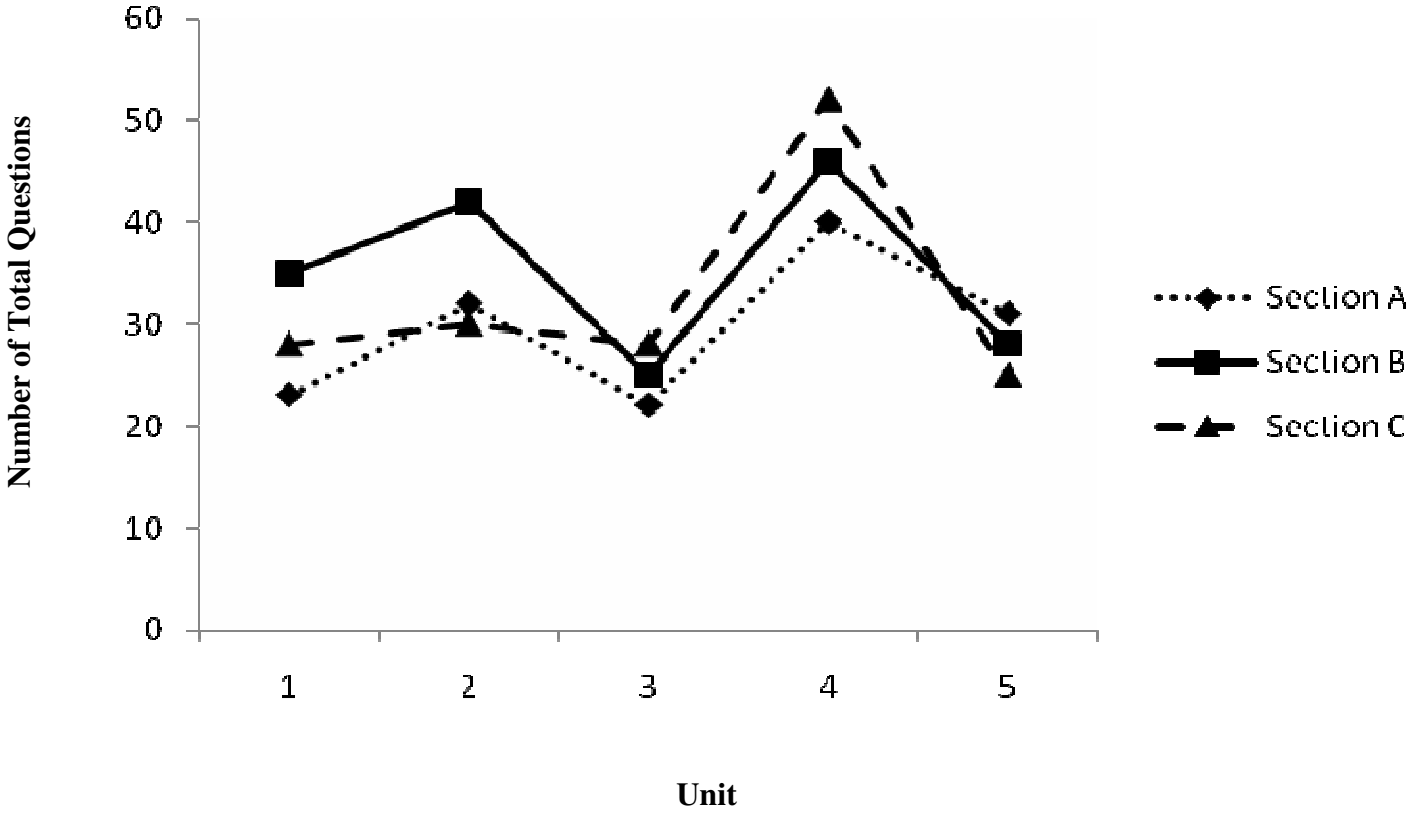


Figure 7. Mean number of total questions per unit posed by the instructor in each section on inter-rater days.

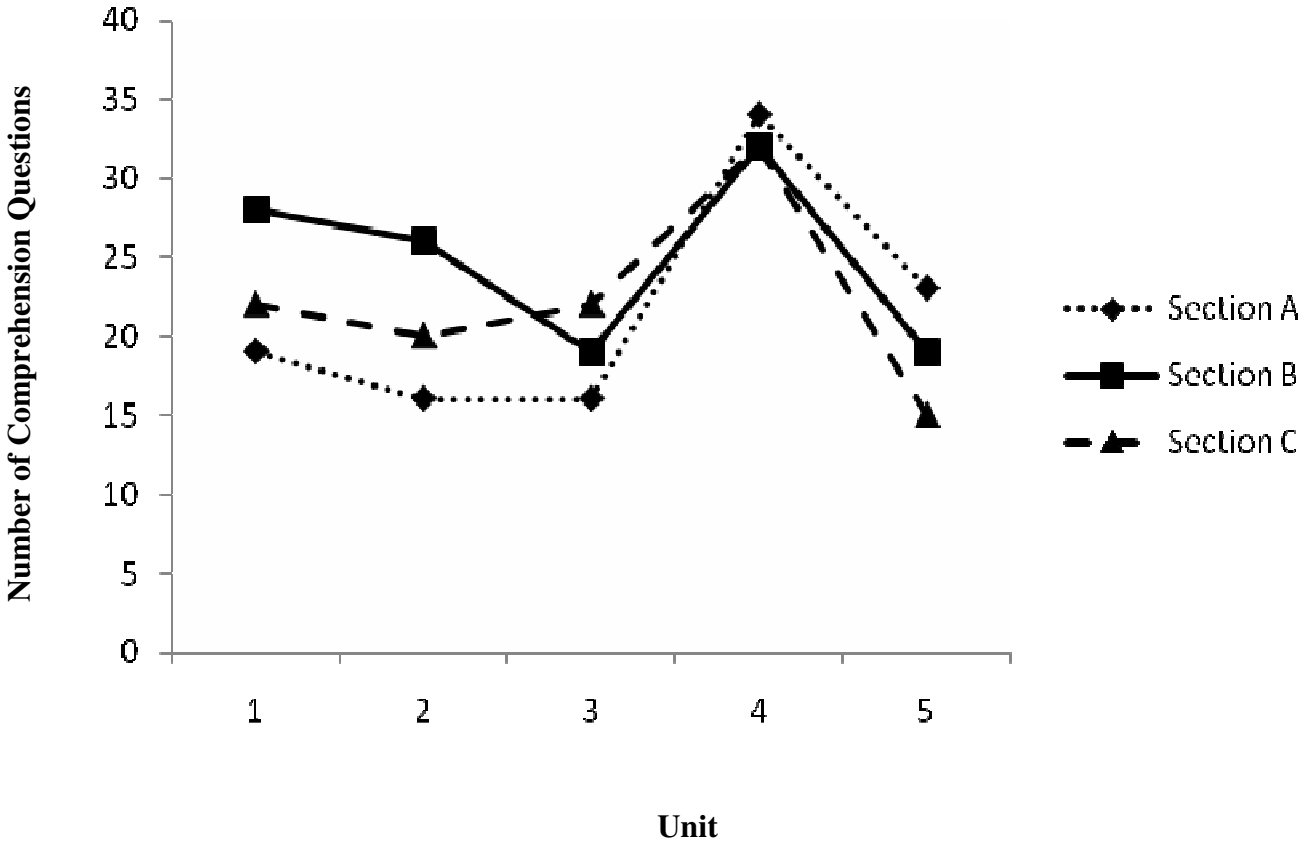


Figure 8. Mean number of comprehension questions on inter-rater check days posed by the instructor in each unit.

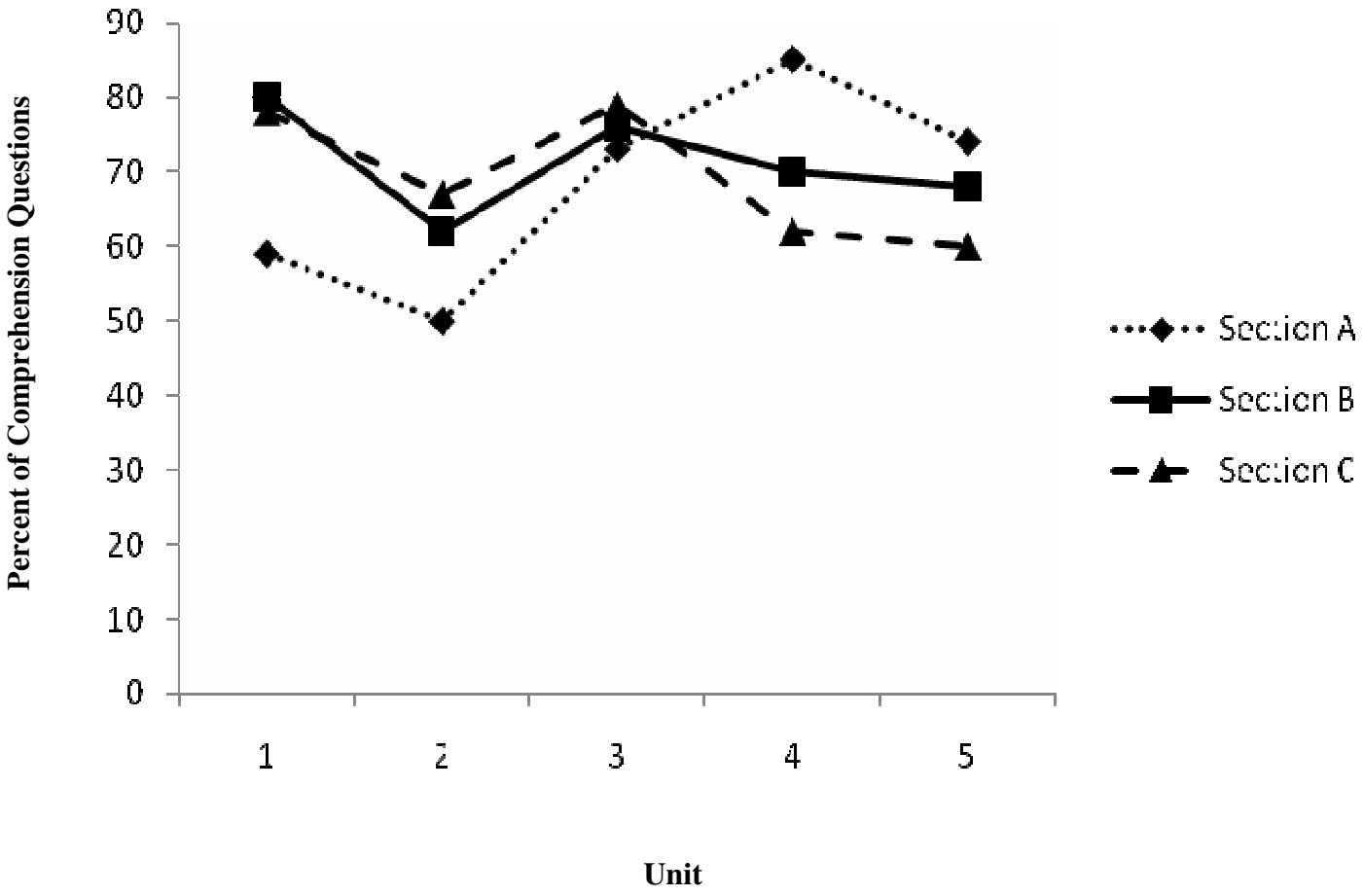


Figure 9. Percent of comprehension questions each inter-rater day posed by the instructor in each section.

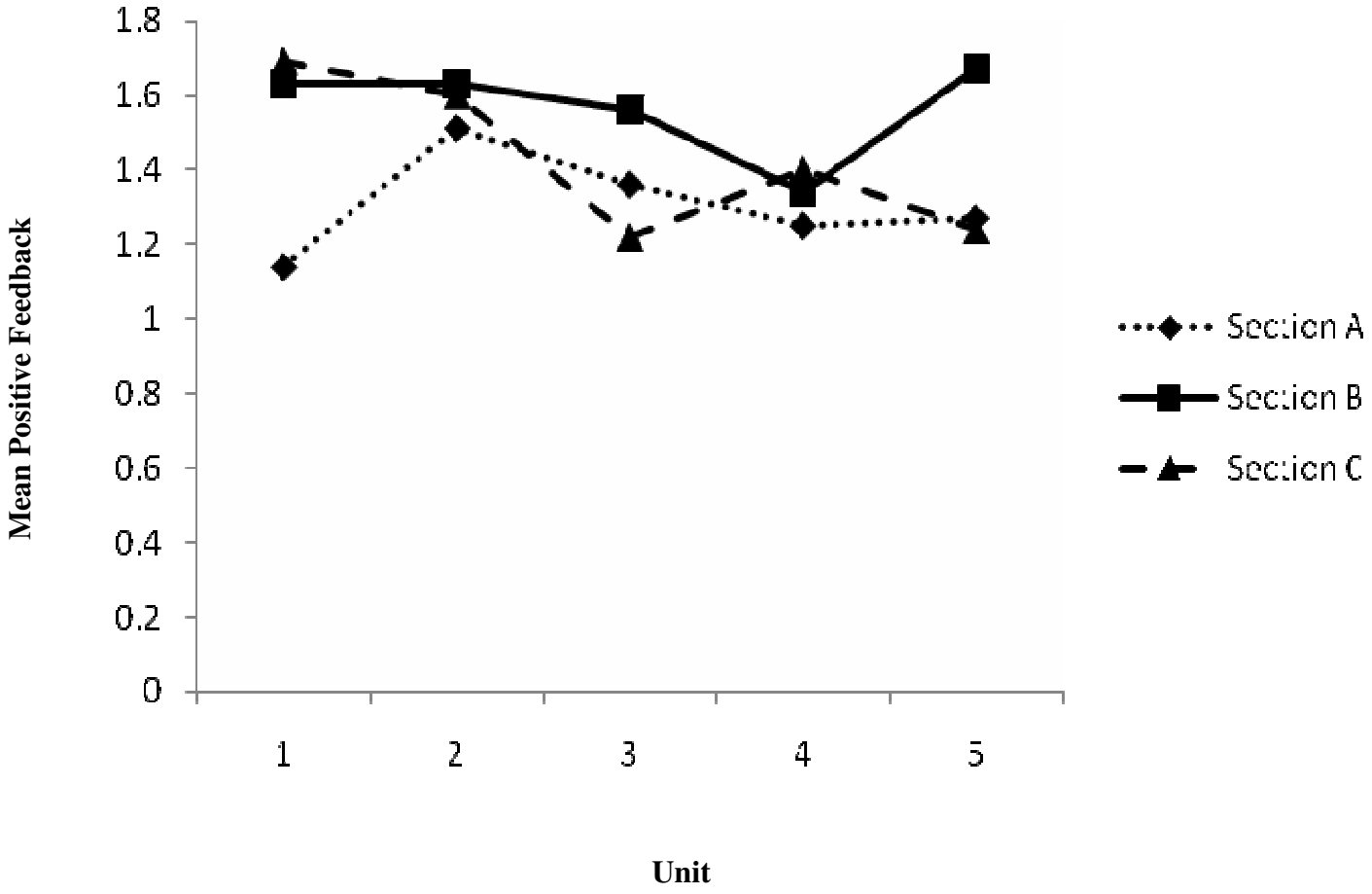


Figure 10. Mean quantity of positive feedback provided to individual students within each unit on inter-rater check days.

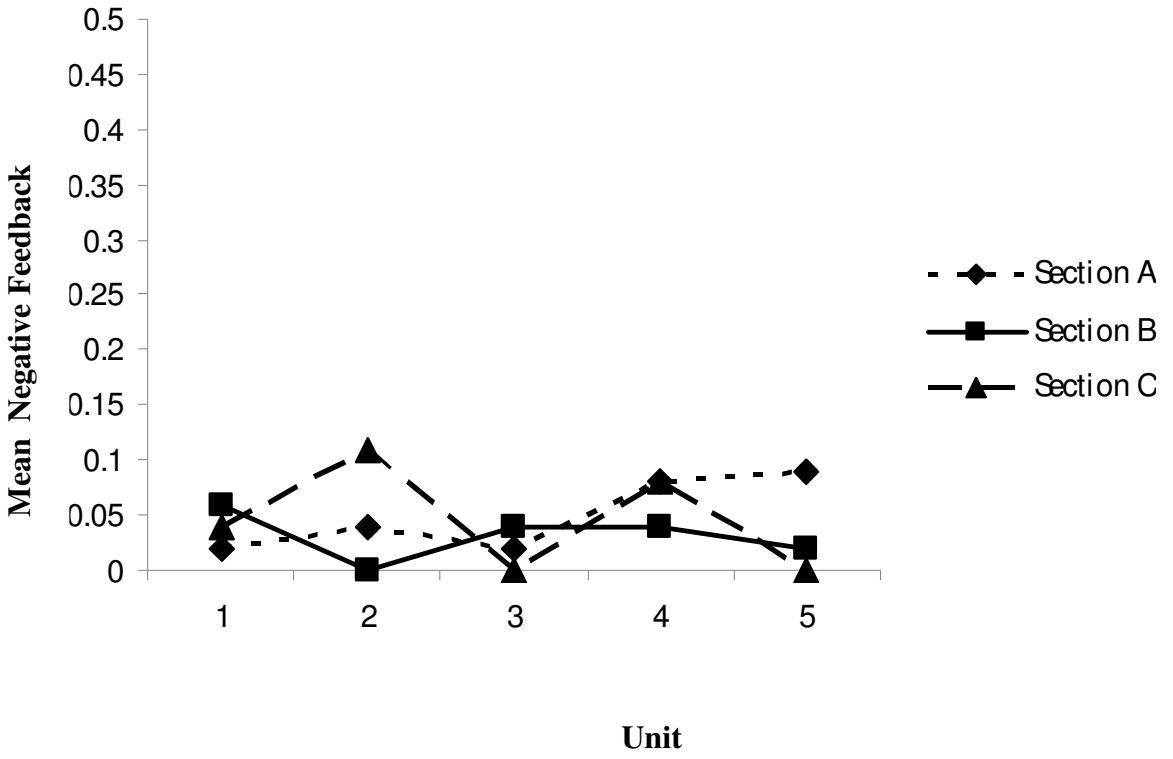


Figure 11. Mean quantity of negative feedback provided to individual students within each unit on inter-rater check days.

Appendix C: Student Discussion Form

Year_____ Semester_____ Date_____ Section_____ Instructor_____ Observer_____

STUDENTS	**T (Timely), R (Repetitious)**											**STUDENTS cont'd**	**T (Timely), R (Repetitious)**									

Appendix D: Instructor Discussion Form

Year_____ Semester_____ Date_____ Section_____ Instructor_____ Observer_____

STUDENTS	FEEDBACK (**+** or -):		**STUDENTS**	FEEDBACK (**+** or **-**):
Factual Questions				
Comprehension Questions				

Appendix E: Students' Under- and Over-Reporting of Class Participation during Non-Credit vs. Credit Units

Student # (Section A)	Non-Credit Units		Credit Units	
	1	2	1	2
1	=	=	=	>
2	=	=	=	=
3	<n	=	>p	=
4	=	>f	=	>f
5	=	=	=	=
6	=	=	=	=
7	=	=	=	=
8	=	=	=	=
9	<	=	=	<
10	<	=	>f	=
11	=	=	<n	=
12	=	=	>	=
13	=	<	<	<
14	=	=	=	=
15	=	=	=	=
16	<	>	>f	=
17	=	=	=	=
18	=	=	=	=
19	<	<	=	=
20	=	=	>p	=
21	=	<	=	=
22	=	=	=	=

Student # (Section A)	Non-Credit Units		Credit Units	
	1	2	1	2
23	=	=	=	>
24	=	<	=	=
25	=	=	=	=
26	=	=	=	=
27	=	=	=	=
28	=	=	=	=
29	=	=	<	=
30	>p	=	=	=
31	<	=	=	=
32	<	=	=	=
33	<	=	=	>f
34	>f	<	=	=
35	=	<	=	=
36	<	=	>f	=
37	>f	=	=	>p
38	=	=	=	=
39	=	=	>f	=
40	=		=	=
41	=		>f	<
42			=	>f
43			=	=
44			>f	=
45			=	

Student # (Section B)	Non-Credit Units		Credit Units	
	1	2	1	2
1	=	=	=	=
2	=	=	=	=
3	=	<	=	<
4	=	=	=	=
5	=	=	=	=
6	=	=	=	=
7	=	=	=	=
8	=	>p	=	=
9	=	=	=	<
10	=	=	=	=
11	<	=	<	=
12	=	<	=	<
13	=	=	=	=
14	<	=	=	=
15	=	=	<	<
16	>	<	=	=
17	<	=	<n	=
18	=	<	=	=
19	=	=	=	=
20	=	=	=	=
21	=	=	=	=
22	=	=	=	=

Student # (Section B)	Non-Credit Units		Credit Units	
	1	2	1	2
23	=	=	=	>
24	=	=	=	=
25	=	=	=	>f
26	=	=	=	=
27	<n	=	=	=
28	=	=	=	=
29	=	=	=	=
30	=	=	=	=
31	=	=	>	=
32	=	=	=	=
33	=	=	=	>
34	=	=	=	=
35	=	<	=	=
36	=	>f	=	=
37	=	=	>	=
38	=	=	=	=
39	=	=	=	=
40	=	>p	<n	
41	=	=	>	
42	<	=	=	
43		=	=	
44		=	=	
45		=	=	
46		=	=	

Student # (Section B)	Non-Credit Units		Credit Units	
	1	2	1	2
47		<	=	
48			=	
49			=	
Student # Section C				
1	<	<	<	=
2	<n	=	<	=
3	=	=	=	=
4	=	=	<p	=
5	=	=	<	=
6	=	=	=	=
7	=	=	=	=
8	=	=	=	=
9	=	=	=	=
10	=	=	=	=
11	=	=	=	=
12	=	=	=	>
13	<	=	>	=
14	=	=	=	=
15	>f	=	=	=
16	<n	=	>f	=
17	=	=	=	=
18	=	=	=	=
19	=	<n	=	=
20	=	=	=	=
21	<	=	=	=

Student # (Section C)	Non-Credit Units		Credit Units	
	1	2	1	2
22	=	=	=	>f
23	<	=	=	=
24	=	=	=	=
25	<n	=	=	=
26	=	=	=	>f
27	=	<	=	=
28	=	<	=	=
29	=	=	=	=
30	=	=	=	=
31	<	=	=	=
32	<	=	<	>f
33	=	<n	=	=
34	<	=	<	=
35	=	=	=	>
36	=	<n	=	=
37	=	=	=	=
38	=		=	=
39	=		=	=
40	=		=	=
41	=		=	>p
42			=	>
43			=	=
44			=	>p
45			>f	=

46				<
47				=
Total =	95	100	111	106
Total >	1	1	5	7
Total >p	1	2	3	3
Total >f	3	2	8	7
Total <	18	15	9	8
Total <n	5	3	3	0

Note. > indicates student over-reporting beyond credit-level; < indicates student under-reporting beyond credit-level; = indicates agreement between student and observer; p = student over-reporting resulted in partial credit being unjustly awarded; f = student over-reporting resulted in full credit being unjustly awarded; n = student under-reporting resulted in students not receiving due credit.

Appendix F: Participation Survey

Following each item stem is the mean and standard deviation for the total sample. The possible range for each item is 1 (option e) to 5 (option a). Note that option a for each item generally indicates a more favorable nature of participation (e.g., greater comfort participating, higher previous or anticipated levels of participation). The percent of students that selected each item option is also provided in parenthesis.

1. Which of the following best describes your previous pattern of class participation in college courses? (M = 3.36, SD = 0.98)
    a. participating several times a day in most class discussions (12.5%)
    b. participating once or twice a day in most class discussions (34.4%)
    c. participating once or twice a day in about 50% of the class discussions (30.6%)
    d. participating infrequently in class discussions (21.9%)
    e. never participating in class discussions (0.6%)

2. Which of the following best describes your expectations for participating in discussions in the 210 course? (M = 3.89, SD = 0.76)
    a. participating several times a day in most discussions (18.8%)
    b. participating once or twice a day in most discussions (55.6%)
    c. participating once or twice a day in about 50% of the discussions (21.9%)
    d. participating infrequently in discussions (3.1%)
    e. never participating in discussions (0.6%)

3. Which of the following best describes your feelings about participating in discussions in classes as large as the 210 course? (M = 3.41, SD = 1.03)
    a. consistently feel comfortable when participating (11.9%)
    b. generally feel comfortable when participating (43.1%)
    c. feel comfortable about half the time when participating (21.3%)
    d. generally feel uncomfortable when participating (21.3%)
    e. consistently feel uncomfortable when participating (2.5%)

4. Which of the following best describes how you feel when called on to participate in class discussion? (M = 3.11, SD = 1.07)
    a. extremely confident (9.4%)
    b. somewhat confident (30.6%)
    c. neutral (25.6%)
    d. somewhat uncomfortable (30.0%)
    e. extremely uncomfortable (4.4%)

5. Which of the following best describes your perspective of the relevance of your comments in class discussion? (M = 4.34, SD = 0.63)
    a. Your comments are almost always relevant. (40.6%)
    b. Your comments are generally relevant. (53.8%)
    c. Your comments are relevant about half the time. (5.0%)
    d. Your comments are seldom relevant. (0.0%)
    e. Your comments are almost never relevant. (0.6%)

6. Which of the following most accurately describes your typical level of preparation for class discussion in past courses? (M = 3.82, SD = 0.53)
    a. completed all of the homework related to the topic to be discussed plus done some additional investigation of the topic (4.4%)
    b. completed all of the homework related to the topic to be discussed (75.0%)
    c. completed most of the homework related to the topic to be discussed (18.8%)
    d. completed little of the homework related to the topic to be discussed (1.9%)
    e. completed none of the homework related to the topic to be discussed (0.0%)

7. What is your attitude about student responsibility for participating in class discussion when the instructor asks students to volunteer responses to instructor questions? (M = 4.10, SD = 0.79)
    a. Students are totally responsible for volunteering comments under those circumstances. (33.1%)
    b. Students bear most of the responsibility for volunteering comments under those circumstances. (46.9%)
    c. Students have marginal responsibility for volunteering comments under those circumstances. 16.9%)
    d. Students have little responsibility for volunteering comments under those circumstances. (3.1%)
    e. Students have no responsibility for volunteering comments under those circumstances. (0.0%)

8. What is your attitude about earning course credit for participating in class discussion? (M = 2.80, SD = 0.90)
    a. Participation should be the most heavily weighted contributor to your grade. (3.1%)
    b. Participation should be substantially weighted in the computation of your grade. (16.9%)
    c. Participation should be moderately weighted in the computation of your grade. (43.1%)
    d. Participation should be minimally weighted in the computation of your grade. (30.6%)
    e. Participation should not be considered in the computation of your grade. (6.3%)

9. How do you feel when an instructor indicates at the beginning of a course that students will be expected to participate in class discussion? (M = 3.12, SD = 0.98)
    a. extremely positive (9.4%)
    b. generally positive (25.0%)
    c. neutral (36.3%)
    d. generally negative (27.5%)
    e. extremely negative (1.9%)

10. How do you feel when an instructor indicates at the beginning of a course that class discussion is welcomed but optional? (M = 4.10, SD = 0.83)
    a. extremely positive (38.1%)
    b. generally positive (35.6%)
    c. neutral (24.4%)
    d. generally negative (1.9%)
    e. extremely negative (0.0%)

11. What would be the relative likelihood of your asking a question versus answering a question in class discussion? (M = 3.32, SD = 1.19)
    a. much more likely to ask a question (20.0%)
    b. somewhat more likely to ask a question (24.4%)
    c. about equally likely to ask or answer a question (31.3%)
    d. somewhat more likely to answer a question (16.9%)
    e. much more likely to answer a question (7.5%)

12. How would you describe the general effect of your participating in discussion on your learning in a course? (M = 3.42, SD = 0.86)
    a. learn best when participating heavily in discussion (6.9%)
    b. learn best when participating regularly in discussion (44.7%)
    c. learn best when participating periodically (33.3%)
    d. learn best when participating infrequently in discussion (13.8%)
    e. learn best when never participating in discussion (1.3%)

13. How do you think that your keeping a record of your participation in class discussion would affect your concentration on the discussion? (M = 3.07, SD = 1.30)
    a. would greatly increase your concentration on the discussion (18.1%)
    b. would generally contribute to your concentration on the discussion (19.4%)
    c. would have an uncertain effect on your concentration on the discussion (27.5%)
    d. would generally detract from your concentration on the discussion (21.3%)
    e. would greatly detract from your concentration on the discussion (13.8%)

14. Which of the following class formats (discussion versus lecture) do you prefer in courses you take? (M = 2.90, SD = 0.97)
    a. all discussion (7.5%)
    b. mainly discussion but some lecture (15.0%)
    c. a balance between discussion and lecture (41.9%)
    d. mainly lecture but some discussion 31.3%)
    e. all lecture (4.4%)

15. What do you see as the relationship between the quantity and quality of class discussion? (M = 2.95, SD = 1.06)
    a. Quantity consistently contributes to quality. (5.6%)
    b. Quantity is more likely to contribute to quality than detract from it. (30.0%)
    c. Quantity and quality are unrelated. (25.0%)
    d. Quantity is more likely to detract from quality than contribute to it. (32.5%)
    e. Quantity consistently detracts from quality. (6.9%)

16. How do you typically respond when an instructor poses a question for class discussion? (M = 3.01, SD = 1.03)
    a. quickly speak up (3.1%)
    b. speak up after a short delay (34.0%)
    c. speak up but with considerable hesitancy (33.3%)
    d. speak up only if no one else speaks up (19.5%)
    e. not speak up even if no one else speaks up (10.1%)

17. How do you feel about a discussion format in which students volunteer comments rather than being called on by the instructor? (M = 4.19, SD = 1.01)
    a. greatly prefer volunteering comments rather than being called on (53.1%)
    b. somewhat prefer volunteering comments rather than being called on (21.9%)
    c. equally comfortable with volunteering and being called on (16.9%)
    d. somewhat prefer being called on rather than volunteering (7.5%)
    e. greatly prefer being called on rather than volunteering (0.6%)

18. Who is responsible for a *high level of student participation* in class discussion? (M = 3.53, SD = 0.74)
    a. exclusively the students (8.1%)
    b. primarily the students (42.5%)
    c. shared equally between the students and the instructor (43.1%)
    d. primarily the instructor (6.3%)
    e. exclusively the instructor (0.0%)

19. Who is responsible for *very limited student participation* in class discussion? (M = 3.24, SD = 0.95)
    a. exclusively the students (17.5%)
    b. primarily the students (25.6%)
    c. shared equally between the students and the instructor (43.8%)
    d. primarily the instructor (17.5%)
    e. exclusively the instructor (2.5%)

20. How would a class with frequent discussion affect your evaluation of the course? (M = 3.41, SD = 1.02)
    a. greatly increase your evaluation of the course (17.5%)
    b. generally increase your evaluation of the course (35.0%)
    c. have little effect on your evaluation of the course (30.6%)
    d. generally decrease your evaluation of the course (14.4%)
    e. greatly decrease your evaluation of the course (2.5%)

21. What effect does frequent discussion by other students have on your concentration in class? (M = 3.52, SD = 0.90)
    a. greatly increases your concentration (13.1%)
    b. moderately increases your concentration (39.4%)
    c. minimally affects your concentration (35.0%)
    d. moderately decreases your concentration (11.3%)
    e. greatly decreases your concentration (1.3%)

22. What effect does your personal participation in class discussion have on your concentration in class? (M = 3.93, SD = 0.90)
    a. greatly increases your concentration (30.2%)
    b. moderately increases your concentration(39.0%)
    c. minimally affects your concentration (25.2%)
    d. moderately decreases your concentration (5.0%)
    e. greatly decreases your concentration (0.6%)

23. What effect does the option of volunteering comments whenever you wish have on your concentration in class? (M = 3.68, SD = 0.82)
    a. greatly increases your concentration (16.9%)
    b. moderately increases your concentration (38.1%)
    c. minimally affects your concentration (41.9%)
    d. moderately decreases your concentration (1.9%)
    e. greatly decreases your concentration (1.3%)

24. How does the possibility that you might be called on to respond to an instructor question affect your concentration in class? (M = 3.77, SD = 1.12)
    a. greatly increases your concentration (28.8%)
    b. moderately increases your concentration (39.4%)
    c. minimally affects your concentration (16.9%)
    d. moderately decreases your concentration (10.0%)
    e. greatly decreases your concentration (5.0%)

25. How does frequent discussion in the class as a whole affect your enjoyment of a class? (M = 3.66, SD = 1.08)
    a. makes the class much more enjoyable (24.4%)
    b. makes the class somewhat more enjoyable (36.9%)
    c. doesn't affect your enjoyment one way or the other(22.5%)
    d. makes the class somewhat less enjoyable (13.1%)
    e. makes the class much less enjoyable (3.1%)

26. How do you feel toward students who frequently comment in class discussion? (M = 3.11, SD = 1.05)
    a. You greatly appreciate their frequent participation. (10.6%)
    b. You generally appreciate their frequent participation. (25.6%)
    c. You feel neutral toward their frequent participation. (31.3%)
    d. You are generally annoyed by their frequent participation. (29.4%)
    e. You are greatly annoyed by their frequent participation. (3.1%)

27. What are your academic expectations of students who frequently participate in class? (M = 4.17, SD = 0.73)
    a. You expect them to do well in the course. (35.6%)
    b. You expect them to do somewhat better than average in the course. (46.3%)
    c. You expect their frequent contributions to be unrelated to their performance in the course. (17.5%)
    d. You expect them to do somewhat worse than average in the course. (0.6%)
    e. You expect them to do poorly in the course. (0.0%)

28. Some students like to be knowledgeable about a course topic before contributing to class discussion on that topic. How do you feel about this issue? (M = 4.58, SD = 0.71)
    a. You have the strongest inclination to comment on topics about which you have the most knowledge. (67.5%)
    b. You are moderately inclined to comment on topics about which you have the most knowledge. (25.6%)
    c. Your knowledge about topics has little effect on your tendency to comment on those topics. (5.0%)
    d. You feel somewhat less need to comment on topics about which you have the most knowledge. (1.3%)
    e. You feel the least need to comment on topics about which you have the most knowledge. (0.6%)

29. To what degree does student sharing of personal experiences in class discussion contribute to the quality of the discussion? (M = 3.71, SD = 0.91)
    a. greatly heightens the quality of class discussion (18.1%)
    b. moderately heightens the quality of class discussion (45.6%)
    c. has a neutral impact on the quality of class discussion (30.6%)
    d. moderately diminishes the quality of class discussion (1.9%)
    e. greatly diminishes the quality of class discussion (3.8%)

30. How do you typically respond when no one else is responding to a teacher question? (M = 3.61, SD = 1.09)
    a. Attempt to answer the question when no one else is responding. (22.5%)
    b. Wait until the silence has become somewhat uncomfortable to you before attempting to answer the question. (40.0%)
    c. Wait until the silence has become extremely uncomfortable to you before attempting to answer the question. (13.8%)
    d. Respond only if the instructor calls on you to answer the question. (23.1%)
    e. Decline to respond to the question even if the instructor calls on you. (0.6%)

31. Do you believe you have insights about course concepts that would benefit your peers if you shared them in class? (M = 3.43, SD = 0.83)
    a. definitely “yes” (10.0%)
    b. generally “yes” (33.1%)
    c. uncertain (48.8%)
    d. generally “no” (6.3%)
    e. definitely “no” (1.9%)

32. What effect do long pauses between teacher questions and student responses have on your desire to participate in class discussion? (M = 3.17, SD = 1.01)
    a. greatly increases your desire to participate (9.4%)
    b. moderately increases your desire to participate (28.1%)
    c. minimally affects your desire to participate (36.9%)
    d. moderately decreases your desire to participate (21.3%)
    e. greatly decreases your desire to participate (4.4%)

33. How would teachers in your past college courses most likely characterize your level of participation in class discussion? (M = 3.00, SD = 0.81)
    a. the most talkative student in class (1.3%)
    b. among the more talkative students in class (26.3%)
    c. talkative to an average level (46.3%)
    d. among the less talkative students in class (23.8%)
    e. the least talkative student in class (2.5%)

34. How would you characterize teacher views regarding the inclusion of class discussion in student grades in your past college courses? (M = 2.89, SD =0.92)
    a. Participation is the most important part of a student's grade. (4.4%)
    b. Participation is among the more important contributors to a student's grade. (21.3%)
    c. Participation is on par with several other contributors to a student's grade. (36.3%)
    d. Participation is among the less important contributors to a student's grade. (35.6%)
    e. Participation is not included in a student's grade. (2.5%)

35. In comparison to other classes you are taking this semester, what expectation do you have for your participation in 210 class discussion? (M = 3.75, SD = 0.79)
    a. more participation in 210 discussion than in any other class (18.1%)
    b. more participation in 210 discussion than in most other classes (41.9%)
    c. about the same level of participation in 210 discussion as in other classes (37.5%)
    d. less participation in 210 discussion than in most other courses (1.9%)
    e. less participation in 210 discussion than in any other course (0.6%)

36. Which of the following best expresses your view of the long-term value of learning to express one's views in public? (M =3.88, SD = 0.77)
    a. Learning to express one's views in public is among the most important skills one can develop in school. (20.6%)
    b. Learning to express one's views in public is among the more important skills one can develop in school. (50.0%)
    c. Learning to express one's views in public is an important skill but certainly not among the more important skills one can develop in school. (26.9%)
    d. Learning to express one's views in public is among the lesser skills one can develop in school. (1.9%)
    e. Learning to express one's view in public is among the least important skills one can develop in school. (0.6%)

37. Your interpretation of how teachers feel about class discussion is best reflected in which of the following claims? (M = 4.30, SD =0.71)
    a. Most teachers strongly value class discussion. (43.1%)
    b. Most teachers moderately value class discussion. (45.0%)
    c. Most teachers are neutral toward class discussion. (10.6%)
    d. Most teachers moderately devalue class discussion. (1.3%)
    e. Most teachers strongly devalue class discussion. (0.0%)

38. How would most of your high school teachers likely describe your participation in class? (M = 3.75, SD = 1.11)
    a. extremely verbal in class (28.8%)
    b. generally verbal in class (36.3%)
    c. occasionally verbal in class (20.0%)
    d. generally quiet in class (11.3%)
    e. extremely quiet in class (3.8%)

39. Many teachers try to stimulate class discussion by asking questions. Which of the following best expresses your view of most teacher questions? (M = 3.55, SD = 0.67)
    a. Most are highly challenging. (6.9%)
    b. Most are moderately challenging. (44.4%)
    c. Most are routine in nature. (45.6%)
    d. Most provide little challenge. (3.1%)
    e. Few provide any challenge at all. (0.0%)

40. At the completion of a class session in which you participated frequently, how would you most likely feel about possible classmate reaction to your comments? (M = 3.41, SD = 0.66)
    a. You would feel your classmates strongly valued your comments. (5.7%)
    b. You would feel your classmates moderately valued your comments. (33.3%)
    c. You would feel that your classmates were neutral toward your comments. (57.2%)
    d. You would feel that your classmates moderately devalued your comments. (3.8%)
    e. You would feel that your classmates strongly devalued your comments. (0.0%)

41. Which of the following best represents your ability to judge the relevance of your comments in class discussion? (M = 4.26, SD = 0.88)
    a. You can determine whether a comment will be relevant even before you make the comment. (50.3%)
    b. You have your first sense of whether a comment is relevant as you are making the comment.(29.6%)
    c. You can tell whether a comment is relevant only by the instructor's reaction to the comment. (15.7%)
    d. You can only judge the relevance of your comment when you have time to reflect on it after class. (4.4%)
    e. You never really have a sense of whether your comment was relevant. (0.0%)

42. How would frequent participation in college classes likely affect your grades in those courses? (M = 3.94, SD = 0.70)
    a. Consistently raise your grades. (18.4%)
    b. Generally raise your grades. (60.1%)
    c. Have little effect on your grades. (19.6%)
    d. Generally lower your grades. (1.3%)
    e. Consistently lower your grades. (0.6%)

43. How do you typically feel when you have volunteered a comment in class discussion? (M = 3.46, SD = 0.90)
    a. You feel very important in the class. (8.8%)
    b. You feel as if you have gained some positive recognition. (46.5%)
    c. You feel neutral about your comment. (27.0%)
    d. You fear that you might have said the wrong thing. (17.0%)
    e. You believe your comment has been poorly received. (0.6%)

44. How much of a personal priority is improving the amount and/or quality of your participation in class discussion? (M = 3.23, SD = 0.80)
    a. It is your top priority. (2.5%)
    b. It is among your highest priorities. (34.6%)
    c. You are neutral about the prospect of improving your class participation. (49.1%)
    d. It is among your lowest priorities. (10.7%)
    e. It is a non-priority for you. (3.1%)

45. What is your opinion of the social status of students who participate frequently in class discussion? (M = 3.19, SD = 0.73)
    a. They tend to be the most popular students in class. (6.3%)
    b. They are among the more popular students in class. (18.9%)
    c. Frequent participation has little effect on one's standing with peers. (63.5%)
    d. They are among the less popular students in class. (10.7%)
    e. They tend to be the least popular students in class. (0.6%)

46. What is your opinion of the social status of students who participate little, if at all, in class discussion? (M = 2.91, SD = 0.72)
    a. They are greatly admired for their quietness. (3.8%)
    b. They are generally admired for their quietness. (6.9%)
    c. Their minimal participation has little effect on how peers regard them. (68.6%)
    d. They are generally discredited for their quietness. (17.6%)
    e. They are greatly discredited for their quietness.(3.1%)

47. What is your view of the relationship between the perceived relevance of course content and student inclination to participate in class discussion? (M = 4.47, SD = 0.74)
    a. Students feel the greatest desire to participate in courses they view as highly relevant. (59.5%)
    b. Students feel a moderate desire to participate in courses they view as relevant. (30.4%)
    c. Student inclination to participate is not affected by the perceived relevance of the course content. (8.9%)
    d. Students feel somewhat less need to participate in courses they view as relevant. (0.6%)
    e. Students feel the least need to participate in courses they view as highly relevant. (0.6%)

48. Which of the following best represents how students' participation in class discussion will affect their personal standing with teachers? (M = 4.27, SD = 0.66)
    a. Students who participate frequently usually are the most liked by their teachers. (36.7%)
    b. Students who participate frequently increase their chances of being liked by their teachers. (55.7%)
    c. Frequent participation has little effect on how much teachers like a student. (5.7%)
    d. Frequent participation decreases students' chances of being liked by their teachers. (1.9%)
    e. Students who participate frequently are the least liked by their teachers. (0.0%)

49. Which of the following best represents the effect of a teacher's friendliness on student participation in class discussion? (M = 4.20, SD = 0.57)
    a. Teacher friendliness is the number one contributor to student participation in class discussion. (27.4%)
    b. Teacher friendliness is among the more important contributors to student participation in class discussion. (65.6%)
    c. Teacher friendliness has little to do with student participation in class discussion. (6.4%)
    d. Teacher friendliness is among the less important contributors to student participation in class discussion. (0.6%)
    e. Teacher friendliness is the least important contributor to student participation in class discussion. (0.0%)

50. Compare the effects of teacher friendliness and teacher knowledge of the subject matter in the course on student participation in class discussion. (M = 3.08, SD = 0.91)
    a. Teacher knowledge is a far greater contributor than teacher friendliness to class discussion. (7.6%)
    b. Teacher knowledge is a somewhat stronger contributor than teacher friendliness to class discussion. (19.7%)
    c. Teacher knowledge and teacher friendliness have an equal impact on class discussion. (49.0%)
    d. Teacher friendliness is a somewhat stronger contributor than teacher knowledge to class discussion. (20.4%)
    e. Teacher friendliness is a far greater contributor than teacher knowledge to class discussion. (3.2%)

Appendix G: Means for Low, Medium, and High-Responding Participants (based on Unit 1 participation levels) on each Item on the Participation Survey

Items on Factor 1-	**Low**	**Medium**	**High**
Expectation for Discussion in College Classes:			
2. Which of the following best describes your expectations for participating in discussions in the 210 course?	3.44	3.93	4.13
8. What is your attitude about earning course credit for participating in class discussion?	2.50	2.77	3.16
12. How would you describe the general effect of your participating in discussion on your learning in a course?	2.92	3.40	3.90
14. Which of the following class formats (discussion versus lecture) do you prefer in courses you take?	2.47	2.93	3.19
20. How would a class with frequent discussion affect your evaluation of the course?	3.08	3.52	3.84
34. How would you characterize teacher views regarding the inclusion of class discussion in student grades in your past college courses?	2.61	2.94	3.06
36. Which of the following best express your view of the long-term value of learning to express one's views in public?	3.69	3.84	3.58
42. How would frequent participation in college classes likely affect your grades in those courses?	3.72	3.95	4.16

Items on Factor 2-	**Low**	**Medium**	**High**
Personal Benefits of Participation:			
21. What effect does frequent discussion by other students have on your concentration in class?	3.42	3.51	3.65
22. What effect does your personal participation in class discussion have on your concentration in class?	3.34	3.90	4.52
23. What effect does the option of volunteering comments whenever you wish have on your concentration in class?	3.25	3.73	3.94
24. How does the possibility that you might be called on to respond to an instructor question affect your concentration in class?	3.42	3.81	3.97
25. How does frequent discussion in the class as a whole affect your enjoyment of a class?	3.31	3.52	4.23
29. To what degree does student sharing of personal experiences in class discussion contribute to the quality of the discussion?	3.42	3.69	4.10
44. How much of a personal priority is improving the amount and/or quality of your participation in class discussion?	2.94	3.24	3.39
Items on Factor 3-			
Personal History and Confidence Regarding Participation:	**Low**	**Medium**	**High**
1. Which of the following best describes your previous pattern of class participation in college courses?	2.61	3.42	3.87
3. Which of the following best describes your feelings about participating in discussion in classes as large as the 210 course?	2.83	3.41	3.87

4. Which of the following best describes how you feel when called on to participate in class discussion?	2.67	3.08	3.52
9. How do you feel when an instructor indicates at the beginning of a course that students will be expected to participate in class discussion?	2.69	3.10	3.48
16. How do you typically respond when an instructor poses a question for class discussion?	2.29	3.11	3.52
31. Do you believe you have insights about course concepts that would benefit your peers if you shared them in class?	3.17	3.45	3.58
33. How would teachers in your past college courses most likely characterize your level of participation in class discussion?	2.33	3.06	3.55
38. How would most of your high school teachers likely describe your participation in class?	3.11	3.90	4.23

All other items:	**Low**	**Medium**	**High**
5. Which of the following best describes your perspective of the relevance of your comments in class discussion?	4.14	4.30	4.61
6. Which of the following most accurately describes your typical level of preparation for class discussion in past courses?	3.94	3.80	3.77
7. What is your attitude about student responsibility for participating in class discussion when the instructor asks students to volunteer responses to instructor questions?	4.03	4.10	4.19

10. How do you feel when an instructor indicates at the beginning of a course that class discussion is welcomed but optional?	4.36	4.16	3.81
11. What would be the relative likelihood of your asking a question versus answering a question in class discussion?	3.50	3.33	3.16
13. How do you think that your keeping a record of your participation in class discussion would affect your concentration on the discussion?	2.92	3.04	3.10
15. What do you see as the relationship between the quantity and quality of class discussion?	3.03	2.89	3.00
17. How do you feel about a discussion format in which students volunteer comments rather than being called on by the instructor?	4.08	4.23	4.32
18. Who is responsible for a high level of student participation in class discussion?	3.44	3.51	3.61
19. Who is responsible for very limited student participation in class discussion?	3.42	3.12	3.32
26. How do you feel toward students who frequently comment in class discussion?	3.14	3.07	3.10
27. What are your academic expectations of students who frequently participate in class?	4.00	4.11	4.45

28. Some students like to be knowledgeable about a course topic before contributing to class discussion on that topic. How do you feel about this issue?	4.56	4.51	4.81
30. How do you typically respond when no one else is responding to a teacher question?	2.97	3.55	4.19
32. What effect do long pauses between teacher questions and student responses have on your desire to participate in class discussion?	3.11	3.04	3.35
35. In comparison to other classes you are taking this semester, what expectation do you have for your participation in 210 class discussion?	3.69	3.84	3.58
37. Your interpretation of how teachers feel about class discussion is best reflected in which of the following claims?	4.22	4.31	4.39
39. Many teachers try to stimulate class discussion by asking questions. Which of the following best expresses your view of most teacher questions?	3.47	3.52	3.77
40. At the completion of a class session in which you participated frequently, how would you most likely feel about possible classmate reaction to your comments?	3.28	3.45	3.45
41. Which of the following best represents your ability to judge the relevance of your comments in class discussion?	4.28	4.22	4.35

43. How do you typically feel when you have volunteered a comment in class discussion?	3.08	3.49	3.77
45. What is your opinion of the social status of students who participate frequently in class discussion?	3.19	3.27	3.13
46. What is your opinion of the social status of students who participate little, if at all, in class discussion?	2.83	2.95	2.87
47. What is your view of the relationship between the perceived relevance of course content and student inclination to participate in class discussion?	4.14	4.47	4.77
48. Which of the following best represents how students' participation in class discussion will affect their personal standing with teachers?	4.33	4.21	4.48
49. Which of the following best represents the effect of a teacher's friendliness on student participation in class discussion?	4.00	4.26	4.35
50. Compare the effects of teacher friendliness and teacher knowledge of the subject matter in the course on student participation in class discussion.	2.81	3.14	3.16

Vita

Daniel Fox McCleary was raised in San Antonio, Texas. In 2006, he received a B.A. in Psychology with a minor in Global Studies from Drury University in Springfield, Missouri. Upon graduating he enrolled at the University of Tennessee to pursue a Ph.D. in School Psychology. In December 2009 he received a M.S. in Applied Educational Psychology. He will receive his Ph.D. in December 2011 after completing a yearlong internship accredited by the American Psychological Association.