

To the Graduate Council:

I am submitting herewith a dissertation written by Katherine Rae Krohn entitled “The Effect of Self-Recording and Contingent Credit on the Quantity and Relevance of College Student Participation in Class Discussion.” 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.)

The Effect of Self-Recording and Contingent Credit on the Quantity and Relevance of College
Student Participation in Class Discussion

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

Katherine Rae Krohn
AUGUST 2010

Copyright © 2010 by Katherine Rae Krohn

All rights reserved.

Acknowledgments

Many people provided support and contributed to the successful completion of this dissertation. It is the product of a creative and hardworking team, whose members are currently conducting a series of related studies. At the heart of this team is Dr. Robert Williams, to whom this dissertation is dedicated. With exceptional creativity, organization, and attention to detail, I have yet to meet an individual who is more committed to collaborating with students to produce high-quality scholarly work than he. I continue to admire the care and discipline of his work. Fortunately, Dr. Williams took time to improve the tone, efficiency, and overall quality of my writing. Even more remarkable is that he instilled in me a love for college teaching. The knowledge and confidence I gained under his mentorship led to a strong interest in evaluating my own instructional practices, which ultimately alleviated the laborious nature of completing a dissertation. I was honored to witness his enthusiasm for this project and have his unwavering support throughout my entire graduate school experience.

It was a privilege to have three of the most prolific and accomplished faculty I know on my committee. I would like to thank Drs. Christopher Skinner, Steve McCallum, and David Cihak for their astute and unique contributions. Their strong knowledge of research design and statistics, as well as willingness to explore implications of my findings and discuss areas for future research was notable. Like Dr. Williams, these experienced professors enjoy high quality discussions and frequent participation in their own classes.

I am indebted to the remaining research/teaching team members. Only good friends and strong students would dedicate time during their busy graduate school education to assist with my arduous data collection procedures. First, I must thank Kathleen Aspiranti for observing three

classes for an entire semester by my side. Beyond her time commitment, her insightful comments and constructive criticism were indispensable. Lisa Foster and Daniel McCleary also devoted substantial time and effort, as well as assisted with the planning and testing of the coding scheme and observation procedures. I wish these students the best and am confident their dissertations will be outstanding. Additionally, I am thankful to the instructors and GTAs in the three classes within which the study was conducted. Cora Taylor, Colin Quillivan, and Meagan Nalls graciously accepted my feedback and adapted their teaching style to accommodate the needs of the study. These instructors, along with their GTAs (Jared Yaw, Charles Galyon, and Carolyn Blondin), have great professional promise.

All aforementioned individuals have greatly advanced my knowledge of research and statistical procedures. My gratitude also goes out to Dr. Jennifer Morrow and Mike Oneil for statistical consulting and the detailed explanations provided.

Finally, I owe a very special thank you to my family. They provided a caring and supportive environment within which I developed a profound love for learning and an inquisitive nature. I am incredibly grateful for their love and encouragement. Without my amazing friends and family, completion of graduate school and this dissertation would not have been nearly as enjoyable.

Abstract

In sizeable discussion-based college classrooms, achieving any approximation of balanced student participation is difficult. More common is a pattern that develops wherein a small percentage of the class dominates discussion and a larger percentage rarely or never participates. Thus, the purpose of this study was to find ways to balance the amount of discussion across students without diminishing the relevance of discussion. Consequently, this study evaluated the efficacy of crediting participation and requiring students to self-record their daily participation. Students ($N = 160$) in three sections of an undergraduate educational psychology course self-recorded their comments on specially designed record cards and received credit for participation during selected phases of the study. Additionally, an observer kept track of each class discussion by coding the quantity of each student's daily participation. Relevance and type of student responses were assessed as ancillary dependent measures, also recorded by the observer.

Credit decreased the percentage of both non-participants and dominant participants, thus balancing participation across students. Self-recording had a minimal effect on participation. Neither credit nor self-recording altered relevance or type of student comments. Few overall instances of non-relevant student commenting indicated that the construct was too narrowly defined, which provides direction for future attempts to assess quality of student participation. Because these findings resulted from comparisons within and between three sections of the course, instructor behavior was also monitored daily. A secondary observer's records revealed that instructor behaviors (i.e., type and number of questions asked and feedback given) did not inflate or diminish the effect of treatment conditions.

A 50-item survey assessed student perceptions of participation at the beginning of the course and was found to significantly predict student participation. Through a series of four principal components analyses, I extracted three specific factors. Logistic regression analyses showed that the primary factor, History and Confidence regarding Participation, differentiated high and low participants as well as the total survey and the three combined factors. This factor best predicted membership in a low participant group in the non-credit units and membership in a high participant group in the credit units.

Table of Contents

Chapter I.....	1
Introduction and Literature Review	1
Student Variables Affecting Participation	2
Course Variables Affecting Student Participation	4
Crediting Student Participation	7
Assessing Student Participation	10
Predicting Student Participation	13
Framework for the Current Study	14
Research Questions	16
Chapter II	18
Method	18
Participants.....	18
Environmental Context for the Study	19
Procedures for Self-Recording Participation	21
Observational Arrangements.....	23
Participation Crediting Procedures	28
Survey of Student Participation Perspectives	28
Participation Contingencies and Research Design.....	29
Chapter III.....	31
Results.....	31
Descriptive Demographic Data and Correlations	32
Reliability Assessments	32
Effect of Credit Treatment on Levels of Student Participation	38
Effect of Self-Recording Treatment on Levels of Student Participation	43
Low-Participating Students and Non-Participants.....	45
Medium- and High-Participating Students	48
Type of Student Participation	49
Relevance of Student Comments.....	51
Interaction between Treatment and Instructor Behavior.....	51
Survey Results	57
Chapter IV.....	68
Discussion & Conclusions	68
Effect of Credit on Patterns of Student Participation.....	69
Effect of Credit on Low-Participating Students.....	70
Effect of Credit on Medium-Participating Students	73
Effect of Credit on High- and Dominating-Participants	74
Reliability and Effect of Student Self-Recording	75
Non-effect of Differential Instructor Behavior	78
Effect of Credit on Relevance of Participation	79
Predictors and Perceptions of Participation	83
Implications for Instructors.....	91
Limitations	93

Future Directions and Conclusions	96
References	98
Appendixes	105
Appendix A: Tables	106
Appendix B: Figures	132
Appendix C: Instructor Discussion Form	144
Appendix D: Student Discussion Form	145
Appendix E: Participation Survey	146
Appendix F: Students' Under- and Over-Reporting of Class Participation during Non-Credit vs. Credit Units	158
Appendix G: Survey Items Comprising Factors, Factor Loadings, and	165
Means for Low, Medium, and High Participants.....	165
Vita.....	172

List of Tables

Table 1. Flowchart of Treatment Combinations and Sequences for Each Section.....	107
Table 2. Mean Daily Participation Level and Standard Deviation Based on Student Gender and Academic Classification	108
Table 3. Correlations between Student and Primary Observer Records of Class Participation on Baseline (b), Credit (c), and Non-Credit (nc) Units.....	109
Table 4. Correlations between Student and Observer Records of Class Participation on Inter-rater Check Days	110
Table 5. Means and Standard Deviations for Student and Observer Records of Class Participation on the Inter-rater Check Day in Each Unit	111
Table 6. Correlations between Student and Primary Observer Records of Relevant Student Participation on Baseline (b), Credit (c), and Non-Credit (nc) Units	112
Table 7. Correlations between Student and Observer Records of Relevant Student Participation on Inter-rater Check Days.....	113
Table 8. Percent Agreement between Student and Observer Records of Non-Relevant Student Commenting on Baseline (b), Credit (c), and Non-Credit (nc) Units.....	114
Table 9. Percentage Agreement between Observer Records of Frequency of Instructor Questions on the Inter-rater Check Day in Each Unit	115
Table 10. Correlations between Primary and Secondary Observer Records of Instructor Feedback to Students on the Inter-rater Check Day in Each Unit	116

Table 11. Differences in Proportions of Students at Different Participation Levels between Treatment Conditions Overall and between Pairs of Adjacent Treatment Units in Each Section	117
Table 12. Number of Low Participating Students in Unit 1 Who Fell into Low, Medium, and High Categories in Subsequent Units	118
Table 13. Number of Medium Participating Students in Unit 1 Who Fell into Low, Medium, and High Categories in Subsequent Units	119
Table 14. Number of High Participating Students in Unit 1 Who Fell into Low, Medium, and High Categories in Subsequent Units	120
Table 15. Mean Number of Specific Categories of Student Participation (Comments, Questions, Anecdotes) per Unit in Each Section	121
Table 16. Number of Relevant and Non-Relevant Comments per Unit Recorded by Students and the Observer	122
Table 17. Number of Questions (Factual, Comprehension, Total) Posed by Instructors by Unit within Each Section	123
Table 18. Proportion of Positive Feedback to Relevant Student Comments and Proportion of Negative Feedback to Non-relevant Student Comments in Each Unit.....	124
Table 19. Mean Scores and Standard Deviations on the Participation Survey for Low-, Medium-, and High-Participating Students	125
Table 20. Logistic Regression Results using the Total Participation Survey Items to Predict Placement in Low and High Participation Groups each Unit	126

Table 21. Percentage of Cases Correctly Classified into Low and High Groups each Unit by Section with the with the Total Participation Survey Score for Each Student as the Predictor Variable.....	127
Table 22. Logistic Regression Results using the Participation Survey Factors as Predictors in each Unit	128
Table 23. Percentage of Cases Correctly Classified into Low and High Groups for the Three Survey Factors together as a Model for Predicting Participation	129
Table 24. Logistic Regression Results using the History/Confidence Participation Survey Factor as the Predictor in each Unit	130
Table 25. Percentage of Cases Correctly Classified into Low and High Groups for Only the History/Confidence Survey Factor as a Model for Predicting Participation	131

List of Figures

<i>Figure 1.</i> Record card submitted daily by students to track participation.	133
<i>Figure 2.</i> Percent of non-participants (0 comment level) each day.	134
<i>Figure 3.</i> Percent of occasional participants (1-2 comment level) each day.	135
<i>Figure 4.</i> Percent of frequent participants (3-4 comment level) each day.	136
<i>Figure 5.</i> Percent of dominating participants (5+ comment level) each day.	137
<i>Figure 6.</i> Percent of low participants in Unit 1 who participated in subsequent units.	138
<i>Figure 7.</i> Mean number of total questions per unit posed by the instructor in each section.	139
<i>Figure 8.</i> Mean number of comprehension questions per unit posed by the instructor in each section.	140
<i>Figure 9.</i> Percent of comprehension questions each day posed by the instructor in each section.	141
<i>Figure 10.</i> Mean quantity of positive feedback provided daily to individual students within each unit.	142
<i>Figure 11.</i> Mean quantity of negative feedback provided daily to individual students within each unit.	143

Chapter I

Introduction and Literature Review

Substantial attention has been devoted to determining strategies to enhance the success of active classroom discussions and participation requirements, or at least justify their use.

Inclusion of participation in the vast majority of course syllabi at the college level attests to its perceived value by faculty (Bean & Peterson, 1998; Jones, 2008). Despite instructors' general acknowledgement of the importance of classroom discussion, participation requirements appear to be de-emphasized in medium- or large-sized undergraduate courses. Boniecki and Moore (2003) attribute this tendency to the desire of instructors to avoid "the silence," or the uncomfortable period of time that follows an instructor question during which no students respond. Another undesirable norm that operates regardless of class size, although its effects are likely exacerbated in large classes, is that of consolidation of responsibility. This norm describes the tendency of only a handful of students to account for most of the participation in a class session (Howard, James, & Taylor, 2002). Having a small number of active participants in a class of 50 or more would result in a great imbalance in discussion across students. Unfortunately, seeing that the effect of course size has been found to be the largest significant predictor of student participation, these norms are rather resistant to change in large classes (Howard & Henney, 1998).

Research overwhelmingly supports the positive contributions that increased participation has on students' success inside and outside of the classroom. Specifically, studies that have offered incentives to bolster participation levels have found increases in student attendance, greater student preparation for discussion, and improved demonstration of students' knowledge

and application of the course material (Jones, 2008). Participation requirements foster accountability for preparation, involve more learners, allow for exposure to diverse perspectives, enhance critical thinking skills, and promote the development of communication skills vital for career success (Bean & Peterson, 1998; Garside, 1996; Jones). Voelkl (1995) concluded that class participation has positive effects on course grades. Additionally, Kember and Gow (1994) examined student learning outcomes as a function of teaching and concluded that interactive teaching methods encourage a “deep approach” to learning (as measured by the Biggs Study Process Questionnaire). On the contrary, mode of teaching that is primarily one-way delivery of information may limit meaningful learning.

Student Variables Affecting Participation

The benefits of student participation are abundant, but not all students are equally likely to participate. Research has identified several factors related to student participation in class discussion. Studies examining gender differences offer mixed results (Auster & MacRone, 1994; Fassinger, 1995; Gilson, 1994; Howard & Henney, 1998; Howard et al., 2002). Second, a substantial body of research indicates that non-traditional students (age 25 or older) are significantly more likely to participate in class than are traditional students (Fritschner, 2000; Howard & Henney; Howard et al., 2002; Howard, Short, & Clark, 1996; Weaver & Qi, 2005). The most cited reasons for lack of participation include shyness, a lack of knowledge, and unformulated ideas (Connor-Greene, 2005; Howard et al., 2002; Mainkar, 2008; Weaver & Qi). Other similar factors that may cause students to refrain from contributing to class discussion relate to their level of confidence, interest, and preparation (Fassinger).

Students' perception of responsibility is also a major contributing factor. Howard and Henney (1998) found that most classroom interactions were student-initiated (e.g., asking questions, offering unsolicited opinions about topics under discussion). Unfortunately, only a handful of students may have a sense of responsibility for participating in medium to large classes. On the contrary, the consolidation of responsibility norm would predict that students generally do not assume responsibility, or that only a handful of students do. Nonetheless, this sense of responsibility depends greatly on the particular instructor's establishing an expectation for participation early in a course. Lecture classes convey that the instructor intends to do most of the talking in class sessions and that student comments may even be intrusive to the flow of the lecture.

Cognitive variables may partly account for the variation in participation across students. For example, research has found participation to be related to pre-course critical thinking and course-knowledge (McCleary, Foster, & Williams, 2010). Critical thinking likely affects students' ability to formulate insightful ideas about class-discussion issues. Pre-course knowledge of course content likely contributes to student engagement in discussion about that content. Although cognitive skills are undoubtedly related to participation, the cause-effect relationships between these skills and participation are unclear. Does participation contribute to critical thinking and course knowledge or do critical thinking and course knowledge contribute to participation? Although little research has addressed the cause-effect relationships between cognitive variables and participation, some studies suggest that participation may contribute to the development of critical thinking (Garside, 1996; Simpson & Courtney, 2008). However, we

speculate that the relationship between participation and cognitive variables is probably reciprocal.

As previously noted, several studies have attributed non-participation to shyness or communication anxiety. Non-participants may have a lot to say but emotional factors may impede their saying it. Aitken and Neer (1993) found that 20% of college students could be classified as having communicative apprehension. Although this factor surely does not account for all non-participation, it may be the single most important factor explaining why some students participate minimally. Personal history may also make a difference. If students develop a habit of participation early in their school experience, that tendency may persist across academic levels. In other words, participation becomes a natural part of the learning experience that likely generalizes across time and classes. Participating early and often in class discussion may prevent communicative apprehension from ever developing.

Course Variables Affecting Student Participation

Four major course variables affect the likelihood of class participation: size of class, seating pattern in the class, instructional approach, and credit for participation. Research indicates that distribution of participation across students is negatively related to size of class. Smaller classes enjoy greater participation than larger classes (Fassinger, 1995; Howard & Henney, 1998; Weaver & Qi, 2005). The best classroom setup to encourage student participation is generally thought to be a circular or semi-circular seating arrangement (Marx, Fuhrer, & Hartig, 1999; Rosenfeld, Lambert, & Black, 1985). The assumption is that facing one another creates a more comfortable arrangement for participation than being seated in rows. Burgess and Kaya (2007) reported that having students all face the same direction was less likely to promote

participation. Montello (1988) found that although location of students' seat only slightly affected achievement, it significantly affected participation. Students sitting in the front and center of the room tend to participate at the highest rate. Sometimes the effects of size and seating pattern cannot be separated. For example, a seminar group is likely to be seated around a table and a class of several hundred in an auditorium. Nonetheless, it is reasonable to assume that both size and seating pattern affect participation.

Beyond the physical features of a class, instructional dynamics also make a significant difference in participation. Little empirical data have been collected on instructor behaviors that affect student participation in discussion despite the well-documented importance of such behaviors (Auster & MacRone, 1994). Research has revealed that instructors' communicating that participation is desirable, safe, and beneficial to students and their peers is imperative, especially for nontalkers (students who speak up only occasionally, if at all) (Howard et al., 2002). Instructors who present themselves as authority figures are less likely to experience verbally engaged classrooms than instructors who incorporate humor, address students by their first name, and ask students to address them by their first name (Nunn, 1996; Weaver & Qi, 2005). A class in which the instructor mainly lectures precludes most participation from the students. Occasionally, students may interrupt the lecture to ask a question or pose a question when the instructor periodically asks if students have any questions. On the other hand, when students are assigned material to study prior to class discussion and then are asked probing questions regarding the assigned materials, the rate of participation is likely to be higher and better distributed across students in the class (Foster et al., 2009).

Students report that they participate most in classes in which instructors call on them when they volunteer, give enough time to answer, provide positive feedback, and ask more analytical than factual questions (Auster & MacRone, 1994). The latter issue, complexity of instructor questions posed, is arguably one of the most important dimensions for stimulating a high quality discussion. Unfortunately, Barnes (1983) found that 82% of college instructors' questions are at the lowest cognitive level as measured by Bloom's (1956) taxonomy. Questions that require only recognition or recall of information (in the knowledge category) are not likely to generate stimulating discussion.

Instructor approaches in initiating student participation in discussion are worthy of investigation. Two major formats that have been used to involve students in class discussion are an open discussion in which students volunteer comments and a "cold-calling" arrangement in which an instructor calls on students either randomly or systematically, targeting particular students (e.g., those not listening, those likely to know the answer). Bean and Peterson (1998) contend that an instructor inclined toward open-discussion likely welcomes any student comment, whereas the cold-calling instructor is usually probing for depth or quality of the student's response. Both approaches have advantages and disadvantages. The open-discussion format allows students to set their own schedule of commenting without fear of being put on the spot. Also, students who volunteer comments may be more likely to feel their comments are class worthy than a student singled out in cold-calling. On the other hand, cold-calling does not allow any student to sit on the sidelines, which presumably would prompt students to pay close attention lest they be called on to contribute. In addition, cold-calling allows the instructor to balance discussion across students.

In addition to posing questions for student responses, instructor feedback (positive vs. negative) could be expected to have an effect on rate and distribution of discussion. Hancock (2000) has shown that verbal praise administered contingently on the target behavior can be very effective with college students. Although negligible research was identified regarding the specific effect of contingent praise on participation at the college level, instructors who validate the importance of student responses by restating a student's answer and/or indicating the importance or correctness of the student's perspective are likely to promote a stronger continuation of responses than those who indicate that the response is incorrect or unrelated to the question posed. Having much the same effect as negative feedback, instructors who ignore student comments are likely to incur fewer student comments in the future. This scenario is most likely to occur when an instructor moves on to another student or issue without commenting on the current student's contribution.

Crediting Student Participation

Perhaps the most direct intervention for increasing class discussion is to give students credit for participating. Of all the performance measures instructors choose to grade, participation is among the most subjective, causing some educators to recommend that it not be graded (Davis, 1993). Jacobs and Chase (1992) criticized the practice of grading participation, arguing that student personality is a major determining factor of participation and shy or introverted students are penalized, record keeping can be problematic, and instructors may have difficulty in justifying assigned participation grades without objective grading criteria. The latter vulnerability is not surprising, considering that many instructors opt to subjectively assign credit at the end of courses without the aid of systematic participation records.

Howard and Henney (1998) proposed that a viable solution to justifying assigned grades is having students grade themselves at the end of class. Foster et al. (2009) and Krohn et al. (2010) used a relatively objective system in which students self-recorded their specific daily comments, with such an arrangement addressing concerns about record keeping and accurate assignment of credit. Smith (1992), in class sizes of approximately 70 students, also instructed students to make and submit notes of their daily class participation.

To address the complaint of student shyness posed by Jacobs and Chase (1992), a primary goal of the current study was to use the incentive of earning course credit to balance student participation across students and assist nontalkers in becoming more expressive in their communication skills. If the expectation of regular student participation is established, awarding credit can be reframed as benefiting shy students rather than doing them a disservice.

In an effort to achieve a balance in discussion and encourage reticent students to participate, Hodge and Nelson (1991) bypassed grading participation altogether by using differential reinforcement principles. A second purpose of implementing the procedure was to illustrate operant conditioning principles prior to the introduction of that topic in class. The instructor put plus marks underneath a student's name when he or she demonstrated the desired behavior change in participation (either an increase or decrease in participation). The marks next to student names served as the only source of reinforcement (no backup rewards were offered for accrued plus marks). The procedure was effective in increasing and balancing discussion across students. However, some students reacted negatively to not receiving reinforcement as expected (particularly the initially high participants). Additionally, this study was done in a class of 14 students, making the arrangement less feasible or desirable in a much larger class.

Token economies have also been used in promoting participation. Boniecki and Moore (2003) provided students with tokens that could be exchanged for extra credit; they offer alternatives for instructors who do not like to use extra credit (dropping a quiz grade, being excused from an exam). They found that the token economy led to increased levels of student participation. Another result that was regarded by the researchers as “impressive,” was students’ shorter response time to instructor questions during the token economy (less than 1 s on average, measured by latency between instructor question and student response time) than when no tokens were available (6 s on average during baseline). Arguably, this consequence can be viewed as undesirable because students’ impulsive responding is likely counterproductive to producing high quality comments.

Sommer and Sommer (2007) found that students had ambivalent attitudes about receiving credit for participation; although students perceived it as increasing classroom interaction and generally contributing to greater discussion quality, some believed it could stimulate irrelevant comments (what some term “playing for points”). The researchers also found that students most in favor of crediting participation were those already participating. Certainly, the intent of awarding credit for participation is not to increase inconsequential comments or to increase participation of frequent contributors. Instead, the goal would be to increase participation from low-participating students and decrease participation from dominating participants, without compromising the relevance of comments.

Despite some criticisms to the contrary (e.g., stimulating irrelevant thoughts and penalizing shy students, Sommer & Sommer, 2007), formally including credit for participation in the grading scheme is recommended by numerous educators (Bean & Peterson, 1998; Dallimore,

Hertenstein, & Platt, 2004; Howard & Henney, 1998; Mainkar, 2007). Janzow and Eison (1990) claimed that participation may be devalued by students if not incorporated into the formal grading scheme. An important issue regarding credit contingencies is the relative advantages or disadvantages of offering regular credit versus extra credit. Although no studies have compared the effectiveness of the two crediting procedures, Boniecki and Moore (2003) found that extra credit successfully increased participation using a token economy in a large college class. It could be argued that the expectation of student participation will only be taken seriously if formally included in the regular course credit scheme. Furthermore, extra credit may not be a powerful incentive to encourage participation among some groups—especially those initially disinclined to participate.

Finally, determining the optimal amount of credit available for participation can be challenging. Foster et al. (2009) offered 2 points credit per day for participation in one semester and 3 points per day in a comparison semester, both of which was a very small percentage of total credit available in the course. Students could still earn an A in the course without any participation credit. Nonetheless, both contingencies increased the percentage of initially non-participating students who subsequently participated in class discussion, though a substantial percentage of students (nearly 40%) remained non-participants throughout the course. Further research is needed to determine the most effective credit arrangement.

Assessing Student Participation

Several methods of assessing student participation are available, including end-of-course surveys, peer-recording or rating, and self-recording, although some may be better suited than others for use in medium or large classrooms. Making subjective decisions about participation at

the end of a course is a relatively simple and unobtrusive alternative to obtaining frequent ratings from individual students. Unfortunately, students may challenge the accuracy of their assigned participation grades based on either peer or instructor ratings, and instructors may not agree with the appropriateness of students' self-assessment of participation. Howard et al. (2002) found that when students were surveyed regarding their participation in class discussion, their estimates were about twice as high as actual observational records. Burchfield and Sappington (1999) also found that students ranked themselves higher on surveys regarding perceived class participation than did their peers and instructors.

Advocates of peer recording cite high levels of inter-observer agreement and validity (Melvin, 1988), but there are several limitations associated with its use. Ratings given by peers are not well accepted by the ratees (Love, 1981). Peer ratings may be more palatable if raters receive the median peer rating, as done by Melvin. Peer ranking, peer nominations, or peer rating can be used. Mainkar (2008) noted that asking students to rank each other on a daily basis is not practical, and using peer nominations (e.g., ranking only the top 20%) does not allow for discriminations among the large remaining percentage of students. Plus, peer ratings done on a daily basis may be too distracting in classes with many students. Mainkar also cautions that student absences and recency effects are problematic aspects of such procedures.

A meta-analysis by Falchikov and Boud (1989) consolidated some findings about the reliability of student self-assessment: students do not consistently underrate or overrate themselves, students in higher-level courses tend to assess more accurately, and better students tend to underrate themselves while poorer tend to overrate themselves. However, self-assessment as measured in the above meta-analysis was much broader than self-recording. Because self-

recording of specific comments is more objective than general self-ratings, reliability of self-recording is the higher of the two self-assessment procedures. In fact, Krohn et al. (2010) found that the reliability of students' self-recorded comments on credit days (.87) and non-credit days (.85) were large and not significantly different. However, when students did not receive credit for their self-recorded participation, student-observer agreement on the specially designed record cards (75% agreement) was not superior to recording on a plain index card (80% agreement). Also, students were much more likely to under-report than over-report the number of comments made.

Another central issue in the assessment of student participation is determining what to assess: quantity or quality of participation. Several studies have demonstrated that crediting participation results in quantitative increases (Boniecki & Moore, 2003; Krohn et al., 2010; Sommer & Sommer, 2007). Quantity measures of participation can include teacher or student measures, subjective or objective measures, and intentional (i.e., hands raised) or actual participation measures. Mainkar (2008) and others have noted that quantity is much easier to assess than quality, which does not have a universal definition. Gioia (1987) suggests that the main determinant of comment quality is whether it contributes to greater peer understanding of the concepts under discussion. Determining quality based upon that definition would require a considerable degree of subjectivity. The operationalization and assessment of quality is far more challenging than assessing quantity, explaining why virtually no studies have included it as a dependent variable. A good first step in the assessment of quality of participation is to record the relevance of students' comments, ensuring that they are on-topic, not redundant, and accurate. In order to record relevancy of student comments, an observer must be intimately familiar with the

course content. The effect of requiring students to self-record the relevance of their comments remains an issue needing further research.

Predicting Student Participation

Combining all the class variables known to affect participation would allow an instructor to predict with considerable precision which classes would participate more than other classes. However, some students will deviate from the participation patterns of the class as a whole (i.e., some students will participate regularly in classes in which most students do not participate and other students will not participate in classes in which most other students participate). A self-report inventory that would predict these student deviations could give instructors a heads up as to what to expect from these students. Particularly problematic to instructors are students who do not participate under seemingly optimal class circumstances. The reasons for this non-participation must lie within the students' personal participation history and perspectives of their role in participation. Once a reticent student's history and personal perspectives of participation are better known, the instructor can privately address these issues with that student and perhaps find an individualized way to engage the student in discussion.

Previous research does not appear to explicitly address the link between student history or perception of participation and actual rates of participation. However, several studies provide insight into student preferences for and perspectives of participation (Dallimore et al., 2004), student perceptions and expectations of participation (Messineo, Gaither, Bott, & Ritchey, 2007), and reasons for participation and nonparticipation (Howard & Henney, 1998). Results from these studies informed the development of the Participation Survey in the current study, provided a point of comparison for the results, and provided the framework for an examination of

potentially strong predictors of participation (e.g., history of participation and confidence in ability to contribute to class discussion).

Framework for the Current Study

The current study most directly extends from the Foster et al. (2009) and Krohn et al. (2010) studies. Both studies employed the same procedures from which the current study builds, although the research questions emanating from each are unique. The Foster et al. study spanned two separate semesters with slight variation in assessment and crediting procedures. In the first semester, students self recorded their comments made in class on a plain 3 by 5 index card and received one point for each comment made for up to two points. Students were credited for and expected to make no more than two comments due to the relatively large class size and time constraints. Lower than desirable record card submission rates in the first semester led to the development of a comprehensive, specially designed record card on which students indicated their completion of other daily credit-producing activities (i.e., homework, attendance, displaying namecard in class) during the second semester. The latter system not only produced superior record card submission rates, but also facilitated quick and accurate assignment of daily student credit. This comprehensive record card was adopted for use in the Krohn et al. study, and in the current study with slight modifications. The amount of credit provided students was increased from the first to second semester in the Foster et al. study, and was further increased in the current study.

The current study addressed some limitations of the Foster et al. (2009) and Krohn et al. (2010) studies: (a) those studies did not monitor instructor behaviors across sections and units within sections, (b) no attempt was made to assess quality or relevance of student comments, and

(c) the effects of self-recording were not separated from other factors (i.e., credit with self-recording was only compared to non-credit phases that also included self-recording). In the current study, observers monitored instructor questions and feedback to determine whether changes in instructor behavior might function as extraneous variables in affecting participation. Also, space was provided on each record card for students to rate the relevance of each comment reported. The effects of self-recording were separated from those of credit to determine if self-recording per se might affect participation.

Krohn et al. (2010) mainly compared student and observer recording of participation to determine the reliability of students self-recording, and concluded that students are unlikely to over-report their comments. However, the authors cautioned that students might have under-reported their comments in non-credit phases. An observer in the current study recorded student participation each day, ensuring that treatment effects were not overstated. The effects of self-recording were separated from those of credit to determine if self-recording per se might affect participation.

Foster et al. (2009) tracked low-responding students (i.e., the bottom quartile of participants based on quantity of comments at the beginning of the course). The researchers found that the credit contingency (including self-recording) increased the participation of initially low-responding participants. Notwithstanding, it was unclear why a sizeable portion of students remained unresponsive. One possibility is that the amount of credit offered was not substantial enough (only 3-5% of the total course credit). Furthermore, some students mentioned a general dislike for the participation crediting arrangement. In the current study, adding greater incentive to participate (more points) and developing a survey to assess student perceptions of

several dimensions of participation addressed these concerns.

A major intent of the survey was to determine what historical, psychological, and/or sociological factors might help explain why some students in Foster et al. (2009) responded to credit and others did not, remaining low or non-participants even when credit was given for participation. In the current study, these self-reported predictors of students' participation in class discussion were identified through principal components analyses and their combined predictive potential for actual class discussion was compared to that of the inventory as a whole. Finally, the predictive potential of each identified factor was compared to that of other factors and the combined factors. The objective was to identify the smallest number of survey items that would predict participation as well as the survey as a whole. A secondary purpose was to determine the class conditions under which the strongest factor would best predict participation. It may be that self-reported history and attitude toward participation better predict participation under some treatment conditions than other treatment conditions (e.g., credit vs. no credit for participation). We hypothesized that student perspective of participation would better predict participation under non-credit than credit conditions. Receiving credit for participation may override a student's personal history and attitudes regarding participation.

Research Questions

The research questions can be classified as primary and secondary. The primary questions relate to cause-effect inferences regarding the effects of contingent credit and student self-reporting on student participation in class discussion (primarily amount but also type and relevance of participation). The secondary questions relate to the extent to which students' self-

reported history and attitudes regarding participation predict their class participation under various treatment conditions.

Primary Research Questions. The primary questions represent a follow-up of our research team's previous research that examined the effects of credit for self-recorded participation on quantity of student participation. The first question in the current study was whether the combination of credit for participation plus self-recording of participation produces more balanced participation (i.e., a greater percentage of participants) than the combination of no credit for participation plus no student self-recording of participation. The next question related to the separation of the effects of credit for participation from the effects of self-recording participation (i.e., whether credit will have a different effect on student participation when students self-record participation than when they do not self-record participation).

Secondary Research Questions. These questions relate to the predictive potential of students' self-reported history regarding their participation (e.g., the extent to which they have typically participated in class discussion, the factors that have previously affected their participation) and attitudes toward student participation (e.g., teacher expectancy of student participation, legitimacy of awarding course credit for participation). It may be that self-reported history and attitude better predict actual participation under some treatment conditions than other treatment conditions (e.g., credit vs. no credit, self-recording versus no self-recording). Given that credit for participation may neutralize some student reluctance to participate, survey scores may predict participation better under non-credit than credit conditions.

Chapter II

Method

Participants

Students ($N = 160$) in three sections ($n = 52, 54$, and 54 in Sections A, B, and C respectively) of an undergraduate educational psychology course participated in the study during the fall semester of 2008. The course was required for all students applying for the Teacher Preparation Program at a large university in the Southeastern United States. The sample consisted of approximately 74.2% women, 25.8% men, 4.1% first-year, 51.4% sophomore, 32.2% junior, 6.8% senior, and 5.5% graduate students. Students' self-reported mean GPA was 3.21, ranging from 1.67 to 4.00, and their course enrollment for the fall semester averaged 14.90 credit hours.

At the conclusion of the semester, students were grouped according to several participation levels in the course, with special emphasis given to low and high participants. Graphs in the Results chapter show the percentage of students participating at different levels on a daily basis: non-participants (0 comments per class session), occasional participants (1-2 comments per class session), frequent participants (3-4 comments per class session), and dominant participants (5+ comments per class session). This grouping permitted assessment of fluctuation in daily participation, within and between phases of the study. Additionally, students initially participating at the following levels per unit were tracked across units in Tables in the Results chapter: low participants (0-2 comments per unit), medium participants (3-7 comments per unit), and high participants (8+ comments per unit). The *ns* for these groups in Unit 1 (baseline) were the following: low ($n = 67$), medium ($n = 31$), and high ($n = 62$). The same

criteria were applied to the mean unit participation for the total course to be used for survey analyses, producing the following *ns*: low ($n = 48$), medium, ($n = 59$), and high ($n = 53$).

A previous study (Howard et al., 2002) used the criterion of 2 comments, at or above which students were referred to as “talkers” and below which students were referred to as “nontalkers.” The class size in Howard et al. was considerably smaller than the 55 student average in the current study, making the criterion for the “nontalkers” impractical. Instead, the cutoff was altered to include a middle group and lowered to accommodate a desirable level of commenting per student based on the larger class size and 50-minute class-session duration. Inclusion of a middle group served to further differentiate the low and high groups, creating a greater possibility that real differences existed between the groups.

Environmental Context for the Study

The course focused on human development issues, with the content divided into five units: Physical Development, Cognitive Development, Social Development, Psychological Development, and Values Development. Typically, seven class periods were devoted to each unit, and each unit followed the same sequence of activities: a video was shown during the first class; assigned course reading material was discussed during the following four class periods; and the final two classes in each unit consisted of a practice exam with feedback and finally a unit exam. The focus of the current study was on days 2 to 5 in each unit, during which students’ recorded class participation in some sections and units, and observers tracked participation in all sections and units.

Participation was examined in 3 sections of the course that met consecutively on MWF mornings, each for 50 minutes. The sections had different instructors, all of whom were enrolled

in the same doctoral program, trained by the same supervising instructor, and given experience as a graduate teaching assistant in the course the previous year. Furthermore, the sections of the course were identical in content, organization, and materials (e.g., reading material, record cards to track participation and other credit-producing activities); differences in the syllabi only pertained to differential self-recording and credit for participation across sections and units within sections.

The course was structured to maximize the opportunity for students to participate in an informed manner. Because student attendance is the necessary first step to set the stage for participation, students received credit for daily attendance. Regular attendance also provided reticent students a chance to observe peers make valuable contributions to class discussion. To emphasize the expectation of participation, Mainkar (2008) recommended that credit incentives be offered to increase attendance but that greater credit be available for class participation. In the current study, one point of course credit was awarded per day for attendance, one point for displaying a namecard, and two points for completing study questions, amounting to 14% of the total course credit available. Four points per day were available for course participation.

Research findings have also suggested the use of assigned readings and study questions to prepare students for class discussion (Hautau et al., 2006; Weaver & Qi, 2005). Student assigned readings, referred to as instructor notes, were five documents (one for each course unit) compiled and updated each year by the supervising instructor. Information included in the instructor notes came mainly from journal articles, but textbooks and scientific reports also provided the basis for selected information in the notes. Each set of notes was approximately 15 typed pages and bound together into one booklet along with other assigned journal article

readings. Before each class discussion in which participation was tracked, students were assigned to read several instructor notes pages and a limited number of journal articles, as well as answer study questions over the assigned readings.

Procedures for Self-Recording Participation

Students purchased a set of specially designed, comprehensive record cards (see Figure 1¹). They were instructed to self-record their specific, individual contributions in class discussion on the record cards during selected days. In accordance with the recommendation of Foster et al. (2009), the record cards were designed to include other daily credit-producing activities in order to increase the likelihood of student submission of the record cards: attendance (signing name on the record card and turning it in), displaying a namecard, completing homework study questions over the assigned instructor notes pages, and completing study questions over assigned journal articles, with each of the four activities worth 1 point per day.

Students also self-recorded the quantity and relevance of their own comments on the record cards. They were asked to jot down a phrase or sentence to represent each comment immediately after making it in class, lest they forget what they had said. A comment was defined as a voluntary question, response to an instructor question, or perspective on an issue under discussion. If a student comment led to a brief interchange between the student and instructor to clarify the meaning and implication of the comment, that interchange was recorded as only one comment. Students' choral responding to an instructor question was not recorded, although simultaneous responding was uncommon because students were instructed to raise their hands.

¹ All figures are located in Appendix B.

Each comment was numbered on the front of the record card, and space was provided on the back of the record card for additional comments. Next to each comment, students indicated whether it was relevant or non-relevant by checking the respective box. Relevant comments were defined in the syllabus as on-topic and consistent with the content of the course, non-redundant, and accurate, whereas non-relevant comments failed to meet any of the relevance criteria. Students judged relevance and non-relevance based on instructor feedback following the student's comment (e.g., "I understand your view," "you've made a very important point," "someone else asked that question earlier," "I'm not seeing how your point relates to what we are discussing"). The first two examples would be judged relevant and the last two non-relevant.

See Table 1² for a specification of treatment combinations and treatment sequence for each section of the course. In two of the three sections, students self-recorded their comments on four days (days 2-5) during selected units. Section A self-recorded comments in two units; Section B self-recorded comments in four units; and Section C did not self-record comments in any unit. No section self-recorded comments in the baseline unit (see Table 1). During non-self-record units, observers were the only individuals that kept track of discussion. Record cards were submitted to the instructor at the conclusion of each class during units in which students were to self-record participation, and at the beginning of class in units in which students were not to record their participation. The latter arrangement allowed instructors to obtain other course records (i.e., attendance, name card and homework credit) and prevented students from self-recording their comments in units specified for no self-recording. Instructors were careful to remind students about the participation contingency operating in each unit: instructors gave

² All tables are located in Appendix A.

students a quiz over the syllabus at the beginning of the course (which detailed specifics of the participation arrangements), posted the contingency for the upcoming unit at the course web site, and displayed the contingency on the board at the front of the room before each class.

Observational Arrangements

The daily observations of two primary observers served three purposes: (1) checking reliability of student self-recording for both the quantity and relevance dimensions, (2) providing the only available measure of student participation during the units involving no self-recording of participation, and (3) systematically assessing instructor behaviors. Both observers had taught the course multiple times, allowing for judgments of student comment relevance and depth of instructor questioning. One observer recorded student comments, while the other recorded dimensions of instructor behavior. Two secondary raters independently recorded either student or instructor behavior on one day each unit to permit assessment of agreement with the primary observers. Observers had the same instructions as the students regarding what constituted a comment. The observers kept a running tally of all comments by each student during the class session. Observer tallies were subsequently compared with the number of comments on student record cards, and the two observers' records also were compared for agreement. To assess agreement between the students' and observers' records, I correlated the total number of comments recorded individually by students with each observer's total for the same students.

Observers sat in the front of the room where they could easily view students' namecards in order to accurately track and record the discussion. Separate forms were created for the tracking of instructor behavior (see Appendix C) and student behavior (see Appendix D). Both forms had students' first names listed alphabetically to make it easier for the observers to match

namecards and listing of students on the observer forms. Instructors were asked to check with observers before class sessions began to check and correct any namecard visibility problems and to call on each student by name during the discussion. These procedural safeguards were necessary for observers to quickly identify and record each student comment or instructor feedback next to the name of the appropriate student. Identical first names were made salient to observers through bold print on the forms, prompting careful checking of last names provided on both the namecards and record forms. The printing of new observational forms was required for each discussion day.

All four observers studied and practiced rating schemes before the first day of class discussion. I composed a sample script of a typical class discussion from which each observer used his or her assigned form (i.e., instructor or student form) to practice coding responses, and comparisons were subsequently made between the two observers rating either the instructor or student comments. Additionally, all observers practiced recording discussion on the first day of Unit 1 (the class before the first regular discussion period) and in the three evening sections of the course not involved in the study.

Observation of instructor behaviors. Instructor behaviors that could affect student participation were tracked across sections and units, with formative observational feedback given to each instructor by the primary observer. These instructor behaviors included type and frequency of instructor questions, as well as frequency of positive and negative instructor feedback to student comments. The focus of collecting data on instructor behaviors was not to determine the effect of the treatment on such behaviors. Instead, the purpose was to promote

consistency in instructors' use of questions and feedback so as to not contaminate any treatment effects resulting from the crediting or self-recording procedures.

Instructor questions and responses to student comments were recorded on all discussion days using the Instructor Discussion Form in Appendix C. The number of instructor questions was tallied, with each tally mark specifying the type of question (factual or comprehension). Factual questions were questions that required straightforward answers—any question worded in such a way that students could respond and fully answer the question by repeating a portion of the instructor notes. Observers did not make judgments about factual versus comprehension questions based on student responses. Each question was classified before hearing the students' responses. For example, the instructor question, "What gender has better eating habits?" was coded as a factual question, although a student might have responded by speculating as to the reasons for specified gender differences. Factual questions were formally defined by the criteria used to assess the bottom-level *knowledge* skill in Bloom's taxonomy: recall of information; knowledge of dates, events, places; and knowledge of major ideas. Comprehension questions were all other questions posed by the instructor, specifically those intended to invite interpretation, prediction, application, illustration, and comparison of the course content. Thus, the term comprehension as used in this study refers to all the additional levels in Bloom's taxonomy (i.e., comprehension, application, analysis, synthesis, and evaluation). Vague instructor questions or those generally intended to open the floor for discussion were not recorded.

Instructor feedback was also monitored. Specifically, feedback was categorized as positive (affirming the accuracy or importance of a student comment or restating the comment)

or negative (indicating the response was off topic, redundant, or incorrect). Negative feedback was also recorded when no instructor statement followed a student comment.

The primary observer gave feedback to the instructors on the few instances when any of their specific behaviors needed to be increased or decreased to promote consistency across units and instructors. Feedback was more prominent in the first unit and particularly the first couple of days in that unit. Each instructor received feedback for one or two of the following issues: needing to ask a greater percentage of comprehension questions relative to factual questions, ask more questions overall, or give feedback to students after each comment (i.e., avoid unintentional ignoring).

Observation of student participation. The number of student comments, the type of comments (i.e., comment, question, anecdote), and the relevance of comments were recorded on all discussion days using the Student Discussion Form in Appendix D. The purpose of recording student participation was to provide a record independent of students' self-reports. When credit was given in the absence of student self-reported records of participation, that credit was based on the observers' records of student participation. Data collected by observers on students' participation was the primary dependent variable used in all analyses. Thus, student self-records were mainly used to assess student-observer agreement, although they were also used to award course credit whenever available.

Each instance of student participation was recorded by the observer as a comment (C), question (Q), or anecdote (A). These codes were used in place of tallies to provide more information on types of participation, permitting later analysis on the amount and pattern of different forms of participation. A question was defined as all student-initiated requests for

information, with the exception of students' asking for something to be repeated or asking about non-content related issues. Many questions that fit into this category were "Can you explain..." or "Why..." questions, although some questions could be confused as a student comment without careful attention (e.g., "I don't understand..." was coded as a question). The Webster's Dictionary definition of an anecdote (2008) was provided to observers: a usually short narrative of an interesting, amusing, or biographical incident. All other types of student participation were coded as comments. This included rhetorical questions as well as opinions that did not meet the criteria for anecdotes. When students made two comments in a row, the observers had to determine if there was a change in topic—otherwise what occurred was considered an interchange between the student and another person and was recorded as one comment.

Observers left each code unmarked if it met the criteria to be considered as a relevant comment, question, or anecdote. A non-relevant comment was indicated with an X through the C, Q, or A. Relevant comments met three major criteria: accurate, on-topic, and not redundant. Comments that failed to meet any of the relevance criteria were coded as non-relevant. Partially accurate comments were rated as non-relevant, but partially complete comments were rated as relevant when they were accurate but just not fully developed. Off-topic comments (e.g., "What homework assignment is due tomorrow?") during the discussion were not recorded. A comment was not coded as redundant if it elaborated on another student's comment, despite some surface similarity. Students did not distinguish between comments, questions or anecdotes on their record cards; however, as previously mentioned, they specified relevant or non-relevant by checking a box next to each response they recorded.

Participation Crediting Procedures

Credit was given for participation in two of the five units in each section: Units 3 and 5 for Section A, and Units 2 and 4 for both Sections B and C. All possible combinations of the treatment conditions were included. The credit ratio was 3 points for the first comment and 1 point for one additional comment each day. More credit was given for the first comment than for an additional comment because of the presumably greater difficulty of entering the discussion than participating a second time in the discussion. Although students could earn full participation credit with two comments per day, the record cards were designed to communicate that making more than two comments was acceptable. Record cards had space for three comments on the front, and for additional comments on the back. Students had the opportunity to earn 16 credit points per credit unit (4 points on each of the 4 discussion days participation was tracked) and 32 points overall, amounting to approximately 6% of the regular course credit.

Survey of Student Participation Perspectives

A 50-item participation survey (see Appendix E) was posted at the course web site and made available for students to download electronically. Student were provided with scan forms on which to record their responses and instructed to turn them in on the second day of class. Scan forms permitted the electronic entry of item response data for each student. All data were recoded into numerical form (e.g., a = 5, e = 1) and the reverse scoring allowed for higher scores to indicate more desirable perceptions and patterns of participation.

Data from this survey were used to predict level of participation in the course and to provide context for understanding why some students commented very little in class, irrespective of credit incentives for class participation. Specifically, questions were designed to assess issues

such as students' participation histories, their comfort level in participating, the perceived contribution discussion makes to a course, and the perceived student versus teacher responsibility for class discussion. All questions had five response options that were designed to be similar to a Likert scale, although it was deemed preferable to write out all options rather than use a simple continuum (e.g., strongly agree to strongly disagree) in order to retain and clarify the meaning of all options. For example, the response options for an item regarding students' perception of who is responsible for a high level of student participation were: (a) exclusively the students, (b) primarily the students, (c) shared equally between the students and the instructor, (d) primarily the instructor, and (e) exclusively the instructor. It is possible that a student may strongly disagree with the claim that participation is exclusively the students' responsibility, but it would be difficult to know if the student strongly feels it is the instructor's responsibility or strongly feels it is a shared responsibility. A series of principal components analyses was done to determine the smallest number of clusters that would retain the predictive potential of the total survey.

Participation Contingencies and Research Design

Decisions about treatment combinations and sequences were informed both by the research questions and the practical needs of the course (e.g., equivalence of participation credit across the participating sections). These issues included the following considerations: (1) each section had the same number of participation credit units; (2) each section by itself provided a sequence of treatments that permitted cause-effect treatment inferences; (3) comparisons of similar units (same content and temporal position in the course) across sections increased the potential for treatment comparisons beyond that of intra-group comparisons in each section,

particularly for comparisons between Sections B and C; (4) a baseline without credit and self-recording was established for each section, permitting the initial assessment of equivalence across sections; and (5) one section had a baseline spanning Units 1 and 2 to replicate the previously found untreated decrement in discussion from Unit 1 to 2 (Foster et al., 2009).

In order to assess the relative effects of credit versus no credit and self-recording versus no self-recording on participation, all possible combinations of the treatment conditions were included. As previously noted, Table 1 displays the treatment combinations and treatment sequence for each section of the course. A random draw determined which section received what treatment sequence. Section A self-recorded participation for two units, Section B self-recorded participation for all four units following the baseline, and Section C did not self-record participation for any unit. Section A permitted a comparison of baseline conditions (NSR-NC) in Units 1, 2, and 4 with the presumably strongest treatment combination (SR-C) applied in Units 3 and 5. Sections B and C allowed for determination of whether self-recording interacted with credit contingencies in accounting for the effects of credit on student participation. Section B permitted a comparison of credit versus non-credit with self-recording held constant for Units 2-5. Similarly, Section C allowed for comparison of credit versus non-credit with non-self-recording held constant for Units 2-5.

Chapter III

Results

Data were mainly examined on a group basis by comparing sections under all treatment combinations (i.e., self-recording versus no self-recording, credit versus no credit for participation) and interactions between the self-recording and credit conditions, although some small group and single-subject comparisons were made for low-participating students. The order of information is presented as follows: (1) demographic and correlational descriptive data, (2) reliability of student self-recording and inter-observer agreement of student commenting; (3) inter-observer agreement of data representing instructor behavior (i.e., questions asked, feedback provided to students); (4) visual representation of levels of student participation—the main dependent variable; (5) levels of student participation analyzed via statistical procedures (proportions testing) to systematically ascertain significance level of treatment effects; (6) intra-student comparisons tracking initially low-participating students across units and treatment conditions; (7) different categories of student participation, including comments, questions, anecdotes; (8) effects of the treatment combinations on relevance of comments as judged by the observers; (9) instructor behavior and possible influences on or interaction with treatment effects; and (10) analysis of the Participation Survey, including presentation of the factors that emerged from the exploratory factor analysis, the overall predictive potential of the survey using logistic regression analyses, and low-participating students' perceptions of participation as measured by the survey.

Descriptive Demographic Data and Correlations

A non-treatment issue of possible interest examined in this section is the differential participation levels of students depending on their gender and academic classification. A point-biserial correlation between mean participation and gender (-.10), as well as a Pearson correlation between mean participation and GPA (-.02) were not significant. Mean participation was significantly correlated with academic classification ($r = .19, p = .05$), although this correlation is small in magnitude. Table 2 presents the mean daily participation level for students by gender and academic classification. Participation rates were slightly higher for males, and slightly higher for first-year and graduate students; however, no differences proved statistically significant. The small sample size for first-year ($n = 6$) and graduate students ($n = 8$) is notable.

Reliability Assessments

Student participation. Student reliability in reporting comments was assessed each day through comparison of student record cards and observer tallies, with the exception of days within non self-recording units (i.e., self-recording took place in only Units 3 & 5 of Section A, Units 2-5 in Section B, and none in Section C). Table 3 shows the daily agreement between students and the primary observer, measured by Pearson correlations. The lowest correlation between students and the primary observer in any section on a particular day and unit was .86, and the highest was .99, which are generally considered to be strong correlations (Cohen, 1988). The level of student-observer agreement was not affected by the contingency condition in effect.

As previously discussed, one primary observer was present each day to track student participation and another to record instructor behavior. Two additional observers were present on the third day in each of five units to assess inter-observer agreement: a secondary observer

tracking student participation and a secondary observer tracking instructor behavior to provide measures of inter-observer agreement for each set of data. Table 3 also shows that student-observer agreement was not affected by the increased number of observers in the room on the third day in each unit, suggesting student habituation to the observers' presence. However, student reactivity to observer presence cannot be ruled out. Students likely understood the purpose of observers' presence without the need for a direct explanation. The most defensible claim that can be made regarding student self-reporting is that accurate records were obtained when observers were present.

Overall, on the inter-rater check day in each unit, Pearson correlations between observers' tallies of student commenting were high (ranging from .94 to 1.0, with most correlations approximating .99), as was reliability between students and both observers' records (ranging from .91 to .98; see Table 4). Reliability between student and observer records might have been reduced because students' under-reported their participation under high levels (e.g., ceasing to record beyond the second comment each day, which maximized the amount of available course credit). Table 5 shows the mean participation levels as claimed by students and recorded by observers on the inter-rater check day, revealing a stable trend in students' under-reporting. Similar to the Krohn et al. (2010) study, patterns of under- and over-reporting were found to be dependent upon the availability of course credit. Students were more likely to under-report participation in non-credit units and over-report participation in credit units. Although the Krohn et al. differences were nonsignificant, students' under- and over-reporting based on credit contingencies was significant in the current study.

Appendix F provides more detailed information on students' under- and over-reporting tendencies. Matches and discrepancies (less than or greater than) between student and observer records are listed per day in this appendix. Significance of the difference in proportions of under- and over-reporting cases was assessed using a statistical procedure to test two independent proportions (see Ferguson & Takane, 1989, pp. 198-200). This proportion-testing procedure was performed using the STAT function on a TI83 graphing calculator. In non-credit units, between 6% and 17% of students per day under-reported their participation (an average of 5.9 cases per day, $n = 54$ for non-credit units, totaling 47/432 cases), compared to between 4% and 8% of students per day who under-reported their participation in credit units (an average of 7.0 cases, $n = 106$ for credit units, totaling 56/848 cases). The proportion of under-reported cases in non-credit units was significantly higher than in credit units ($p < .01$). As predicted, student under-reporting generally occurred with high levels of student commenting (i.e., when more than 2 comments were made per day). In other words, with the exception of 6 total cases, students under-reported only beyond the credit contingency to ensure they would still earn full participation credit.

Student over-reporting in a non-credit unit was virtually nonexistent (i.e., only one student on one non-credit day). However, student over-reporting did occur in credit units. Between 2% and 7% of students per day in credit units over-reported their participation, with the majority of these cases allowing students to receive partial participation credit that was undeserved. The proportion of over-reporting cases in non-credit units (an average of 0.0 cases, $n = 54$, totaling 1/432 cases) was significantly lower than the proportion in credit units (an average of 4.4 cases, $n = 106$, totaling 35/848 cases; $p < .01$). Only one student on one day received full

participation credit when no credit should have been awarded. Student over-reporting when it had no effect on credit (i.e., student over-reporting beyond the 2-comment criterion) was rare, occurring in only 6 total cases. The proportion of cases involving over-reporting above the contingency level (an average of 0.75 cases, $n = 106$, totaling 6/848 cases), which did not affect their credit, was significantly lower than over-reporting to receive credit (an average of 3.6 cases, $n = 106$, totaling 29/848 cases; $p < .01$).

Within-student comparisons across the eight credit days revealed that students did not greatly abuse the record card system to receive undue credit. Sixteen total students (16% of all students who self-recorded) over-reported their participation to receive partial credit on only one of the eight credit days. More regular over-reporting occurred for only a handful of students: three students over-reported to receive credit on two of the eight days, one student over-reported on three days and one student over-reported on four days. In the main, it would not be fair to say that students intentionally over-reported their comments for credit, especially considering that it only happened on one occasion for most students. Although unlikely, it is possible that the mismatch between student and observer records represents an observer error in the record keeping system. Krohn et al. (2010) also concluded that students seldom claimed participation credit that the observer records showed was not earned.

In addition to reliability for frequency of student commenting, agreement between students and observers was also obtained for relevance ratings of student participation. Similar to the recording procedure for overall participation, observers tallied the number of relevant and non-relevant comments made by each student during discussion class periods and students indicated whether each of their comments was relevant or non-relevant on their specially-

designed record cards. Table 6 displays the daily reliability coefficients between the students and primary observer's relevance ratings. Reliability between students and both observers for relevance ratings, as well as between the two observers on the inter-rater check day in each unit is shown in Table 7. Pearson correlations between students and observers on the inter-rater check days ranged from .84 to .98. Inter-observer agreement for relevant student comments was always .99 or 1.0.

Pearson correlations were not computed for reliability of non-relevant comments. The mean number of non-relevant comments reported by students in each section (Sections A and B) was a mere 8 per unit. In other words, considering each unit spanned a four-day period, non-relevant commenting as reported by students was very infrequent. Percent of agreement was chosen in lieu of standard correlations due to the extremely small *ns*, as such correlations would have been inflated due to the large number of 0-rating agreements. Table 8 displays the percent of agreement between students and the primary observer based on daily records, as well as the percent of agreement between the two observers based on records from the inter-rater check days. On the five inter-rater days (one in each unit), there were 43 non-relevant ratings given by observer 1 and 44 by observer 2 for corresponding students, producing agreement on 98% of the total cases. The one instance of disagreement was in the baseline unit of Section C. Table 8 does not display levels of agreement between students and observer 2 because no new information would be provided (due to the superior agreement between observers and because agreement levels with the second observer are based on only one day per unit).

Despite the high inter-observer agreement, agreement between the students and observers within a section and unit ranged from 0% to 30%. The poor agreement cannot be exclusively

attributed to students reporting fewer non-relevant comments than observers. In several cases observers did not agree with the non-relevant distinctions made by students. For example, for the unit in which the highest agreement level was obtained (30% in Unit 4 of Section B), students and the primary observer agreed on only 8 out of 27 cases, students reported non-relevant comments when the observer did not on 5 of 27 cases, and the observer reported non-relevant comments when students did not on 14 of 27 cases. As another example, the breakdown in the unit with the lowest agreement level (0% in Unit E of Section A) was as follows: 0 out of 16 cases where the observer and students agreed, 5 out of 16 cases in which students reported more non-relevant comments, and 11 out of 16 cases in which the observer reported more non-relevant comments.

Instructor behavior. The instructor behavior data were collected as a control variable, permitting comparisons of the three instructors in regard to their question and feedback to students. Pearson correlations were calculated to assess reliability of instructor feedback. A traditional correlational analysis was not possible for the assessment of instructor questions because the only available data were agreement comparisons based on the 20 discussion days; instructor questions had no direct relationship to particular students, rendering it impossible to enter this data into the comprehensive SPSS database. On the third day in each unit, agreement between the primary and secondary observers for instructor questions was calculated by dividing one rater's frequency count by the higher frequency count for each type of question (factual, comprehension, and total). The percentage agreement ranged from 82% to 100%, and did not differ by contingency, unit, or type of question (see Table 9). Reliability coefficients for instructor feedback varied substantially depending on the type of feedback. Specifically,

correlations were much higher for positive feedback, ranging from .95 to 1.0. There were considerably fewer instances of negative feedback and, therefore, slight disagreements impacted the correlations more than for positive feedback; correlations for the negative feedback ranged from .36 to 1.0. As Table 10 shows, with the exclusion of the .36 correlation, all correlations were significant at the .01 level.

Effect of Credit Treatment on Levels of Student Participation

The effects of credit for participation on students self-reported level of participation were examined in two ways: first, visually presenting the percent of students participating across treatment and non-treatment phases in each section; and secondly, using proportions testing to determine if the proportion of students participating at different levels differed significantly between treatment and non-treatment phases. The purpose of the multiple analyses was to determine whether treatment effects would be consistently strong or weak, with a percentage of students participating at different levels per day constituting the main dependent variable: 0 level, 1-2 level, 3-4 level, 5+ level. Before making these determinations of differential participation levels, I assessed equivalence of students' prior knowledge across the sections, as well as equivalence of students' initial participation levels during baseline.

Equivalence of students' prior knowledge across sections. To determine if students in the three sections could be considered equally knowledgeable about information to be presented in the course, potentially impacting their level of participation, I compared their scores by section on a pre-course exam. The mean scores on the exam were virtually equivalent across the three sections, with each mean near 22 out of 50 items. Specifically, the pre-course mean was 21.25 for Section A, 22.17 for Section B, and 21.65 for Section C. A one-way analysis of variance

revealed these means not to be significantly different, $F(2, 152) = .587, p = .557$. Because of equivalence in pre-course means across sections, this dimension was not included as a covariate in the remaining data analyses. The pattern of pre-course means not only suggested equivalence across the sections but also limited entry-level knowledge regarding concepts and issues in the course. However, within each section, the pre-course scores varied from 12 or 13 to 32.

Although it is feasible that students' prior knowledge could affect their inclination to participate, negligible correlations were found between pre-course scores and participation in most phases of the study, indicating that pre-course knowledge was minimally predictive of student responsiveness to credit or self-recording conditions. Nevertheless, the correlation of .18 between pre-course scores and students' overall participation means was significant at the $p = .05$ level. Closer inspection of the differences in correlations between sections revealed that correlations were only significant in Section B, which was the section implementing the self-recording contingency in Units 2 through 5. Correlations between pre-course knowledge and participation means were notably higher in the self-recording units in Section B (i.e., between .30 and .39) than in the baseline unit (.20). Overall, correlations between pre-course scores and self-recording scores (.36 for the SR-NC condition, .27 for the SR-C condition) were stronger than for credit contingencies (.04 for the NSR-C condition, .15 for the NSR-NC condition). Thus, credit effects tended not to be differentially affected by pre-course knowledge, whereas self-recording effects appear to have reflected some level of pre-course differences.

Equivalence of baseline across sections. A one-way analysis of variance was completed to determine whether significant differences in mean levels of participation were present between sections during the baseline unit (Unit 1). The analysis did not yield a significant effect, $F(2,$

157) = .782, $p = .459$. Another important determination regarding the baseline levels was whether they appeared to be elevated in comparison to participation levels in a subsequent unit when no contingency was operating (i.e., an extended baseline). Previous studies (Foster et al., 2009; Krohn et al., 2010) that set the background for the current study reported artificially high baseline levels of participation, presumably due to either the nature and familiarity of the content covered, the newness of the course, or a combination of the two. This study documents the same effect. The baseline in Section A was extended past the first unit, after which mean participation per day dropped per student from 1.42 in Unit 1 to 1.31 in Unit 2. Interestingly, mean participation levels also dropped from Unit 1 to Unit 2 in the other sections despite the addition of credit contingencies in both Sections B and C and a self-recording contingency in Section B (from 1.86 to 1.56 in Section B, and 1.79 to 1.64 in Section C); note that the contingencies were not predicted to increase participation overall, but rather to better balance discussion across students (i.e., the dependent variable was percent of participants at different levels instead of frequency of participation per student).

Visual analysis of mean participatory levels. Figures 2 through 5 show the percent of students in each section who participated in class discussion at particular levels in the various contingencies. The participation levels of greatest interest in this study include those below the credit contingency (0 comments per day), meeting the credit contingency at least partially (1-2 comments per day), slightly exceeding the credit contingency (3-4 comments per day), and greatly exceeding the credit contingency (5+ comments per day). A visual inspection of the figures shows a consistent trend: more students participated when credit than non-credit contingencies operated. For example, as shown in Figure 2, the average percent of students who

did not participate at all fell from 51% in non-credit units to 27% in credit units (Section A), from 48% to 23% (Section B), and from 49% to 30% (Section C). Stated differently, the percentage of non-participants was cut nearly in half in credit units.

The introduction of the first credit phase was staggered such that Sections B and C received credit for participation after four days of baseline (i.e., at the beginning of Unit 2), whereas Section A remained in baseline until Unit 3. The percent of non-participants decreased with the implementation of each credit unit and increased during subsequent non-credit phases in all sections.

Alternative graphic displays of the treatment effect of credit includes consistent increases in the number of students making 1 or 2 comments, as well as some decreases in the number of students commenting excessively. The percent of students making between 1 and 2 comments increased approximately 26% in credit units (Figure 3). No major changes were evidenced for students at the 3 and 4 comment level as contingencies changed (see Figure 4). Figure 5 illustrates a reduction in the percent of students who tended to dominate class discussion (making 5+ comments per day) by approximately 7% in credit units.

Proportions testing. To assess the statistical significance of differences in proportion of participants from one contingency to the next, as seen in Figures 2 through 5, the statistical proportions testing procedure cited earlier in this chapter was used (Ferguson & Takane, 1989). For each section, comparing the proportion of students at each participation level from all non-credit units (including the baseline unit) with the proportion of students from the two remaining credit units resulted in p values indicating statistical differences between credit and non-credit units. Additionally, pairs of contingencies were compared to generate p values indicating the

change in proportion from one phase to the next phase (i.e., treatment phase followed by a non-treatment phase). All proportions and resulting p values are presented in Table 11. Comparisons in this section focus on the effect of credit irrespective of the self-recording contingency, which will be discussed later.

The proportion of non-participants (0 comment level) was significantly less ($p < .01$) in credit units than in non-credit units for all three sections. All adjacent treatment and non-treatment comparisons in each section were also significant ($p < .05$). Specifically, for Section A, the first pair of contingencies (SR-NC and SR-C) showed a significant decrease of non-participants ($p = .006$) from the non-credit to credit phase, as did the second pair ($p = .008$). For Section B, the first pair of contingencies (SR-C to SR-NC) showed a significant increase of non-participants from the first credit to non-credit phase ($p = .024$), with a greater increase for the second pair ($p = .0009$). For Section C, the first pair of contingencies (NSR-C to NSR-NC) also showed a significant increase of non-participants from the first credit to non-credit phase ($p = .018$), as did the second pair ($p = .005$).

For each section, the proportion of students either partially or fully meeting the credit contingency (1 to 2 comment level) was significantly higher during credit than non-credit units. All combined and adjacent-pair comparisons resulted in p values $< .05$ or beyond (most p values $< .01$). On the contrary, no significant differences were found for any section within the 3 to 4 comment level (above the credit contingency level but not considered extreme or dominating). Proportions of students who commented at the 3 to 4 comment level remained relatively consistent despite changes in the contingency. The greatest variation seen at this level was for an adjacent-pair comparison in Section B (a reduction of an average of 11 students in the credit

phase to 7 in the non-credit phase), and an adjacent-pair comparison in Section C (a reduction of an average of 13 students in the credit phase to 7 in the non-credit phase), although neither difference in proportions proved significant. Percentages of students participating at the 3 to 4 comment level (slightly above the credit contingency) were not predicted to increase or decrease as conditions changed. Instead, credit offered in a condition was predicted to decrease the number of non-participants, increase 1 to 2 level participants, and decrease dominating participants (5+ comments per class period). As already described, the former predictions for the 0 and 1-2 level participants were supported. The latter proved true for two of the sections. For Sections A and B, the proportion of students participating far above the credit contingency (5+ level) was significantly lower during credit than non-credit phases ($p < .05$). However, this pattern did not hold for all adjacent-pair comparisons ($p = .09$ for the second pair in Section A, and $p = .07$ for the second pair in Section B), or for any comparisons in Section C.

In sum, the credit contingency for Section A, which included self-recording, appears to have been very effective compared to a withdrawal contingency involving no credit or self-recording. This finding replicates the conclusion from the Foster et al. (2009) study. The credit contingencies in Sections B and C had a similar, effective result. However, the ordering of contingencies in the latter two sections were designed to isolate any potential impact of self-recording by itself or in combination with the crediting procedures on percentage of students at different participation levels.

Effect of Self-Recording Treatment on Levels of Student Participation

To discern whether the self-recording procedure affected participation, apart from the credit contingency, I compared percentages of participants at different participation levels in

Sections B and C. The only difference in the contingency arrangement between these sections for each unit was whether students self-recorded. In other words, credit contingencies occurred in the same units (Units 2 and 4) in both sections, whereas only students in Section B self-recorded their participation. Specifically, if self-recording had a substantial effect on participation, one would expect the percent of participants to be much greater in Section B than Section C in Units 2 through 5, inasmuch as self-recording was an additional component in the former section. Additionally, one would expect minimal between-unit differences within each section (i.e., between non-credit and credit units in Section B and between non-credit and credit units in Section C) because self-recording was held constant throughout the units in each section. As shown in Figure 2, neither of these trends occurred. (Note that this figure displays the percent of *non-participants* in each unit. Thus, to determine the effect of the treatment on the number of *participants* at any level, the trends would need to be reversed, e.g., 25% non-participants in Unit 2 of Section B can be understood as 75% participants.) Instead, the mean percent of participants was not substantially higher in comparable units for Section B than for Section C (i.e., there were only slightly more participants in Section B than Section C: 75.0% vs. 67.3% in Unit 2, 56.0% vs. 45.5% in Unit 3, 78.4% vs. 73.2% in Unit 4, and 48.1% vs. 48.0% in Unit 5). Proportions testing confirmed that the apparent differences between sections in Unit 2 ($p = .176$) and Unit 3 ($p = .155$) were not significant. Additionally, large differences existed between units within each section. As previously discussed, the latter effect resulted from the credit contingencies (more students participating in Units 2 and 4 in both sections, irrespective of self-recording) and is the predominant visual trend in the figure.

Low-Participating Students and Non-Participants

Analyses thus far have focused on potentially inter-subject comparisons to determine the percent of students participating at the various levels across units and contingencies, as well as across sections. Many of the students participating at low levels (i.e., average participation per unit below 0.5 comments; criterion described below) were actually the same students across units. There were 104 students that could be classified as low participants in at least one unit, which amounts to 65% of the total sample. Almost half (47%) of these students (or 23% of the total sample) were low participants across 4 or all 5 of the units. For the latter group of students, we can assume the credit and self-recording treatments were ineffective in increasing participation rates. Reasons for the lack of a treatment effect with this group are examined in the context of the Survey data and in the discussion chapter. Of additional interest are the students that began the course with low rates of participation. It is important to distinguish the initially low-participating students for which the treatments had an effect from those who remained in the low group across units. The intra-subject comparisons made in this section allowed the tracking of students with initially low rates of participation across units and contingencies.

Low-participating students were classified as those with an average of 0.5 comments or less per day in Unit 1, which approximated the bottom third of the Unit 1 participation distribution. This criterion was chosen for both conceptual and practical reasons. First, the criterion of 0.5 comments or less for the daily average within a unit would mean that a student would comment no more than twice in a given unit (twice in one day or spread throughout the four days), which represents a genuinely low rate of responding. In fact, other researchers (e.g., Howard et al., 2002) have used a more modest criterion of less than 2 comments per class

session to refer to the infrequent participants, labeled as “nontalkers” (although the 2-comment criterion also may be considered a reasonably low rate of responding in the Howard et al. study due to the smaller class size). In the current study, a genuinely high rate of responding was determined to be maximizing credit or exceeding the credit limit (i.e., an average of 2 or more comments per day within a unit). The high group had approximately the same number of students as the low group for each section. In addition to examining the performance of low-participating students, as defined above, a more conservative subgroup was also examined. The majority of the low-participating group made 0 comments in Unit 1 (this subgroup of low-participating students is referred to as non-participants) and their participatory behavior was also tracked across units and conditions. In total, there were 67 low-participating students ($n_s = 23$, 22, 22 in Sections A, B, and C), which includes the non-participant subgroup (43 non-participants: $n_s = 14$, 15, 14).

Figure 6 shows the percent of initially low-participating students who subsequently participated each day. Across units, there were consistent and dramatic increases in the mean percent of low participants who participated from non-credit (with baseline included in the mean) to credit units: from 12% to 46% in Section A, from 18.1% to 52.6% in Section B, and from 9.9% to 41.4% in Section C. All sections show clear increases in each credit unit and subsequent decreases in each non-credit unit, which also dilutes the effect of self-recording on this sample. Proportions testing results provides additional support for a strong credit contingency effect on participation for low-participating students. Overall, the proportion of initially low participating students who participated in credit units (16/22) was significantly higher ($p = .003$) than in non-credit units (7/22) in Section A, significantly higher in credit units

(15/22 vs. 8/22; $p = .017$) in Section B, and significantly higher in credit units (16/22 vs. 5/22; $p = .0005$) in Section C. Adjacent treatment and non-treatment unit pairs were all significantly different as well.

A more precise portrayal of how the credit contingencies affected the low-participating students is shown in Table 12. The overwhelming majority of initially low participants remained in the low group for subsequent non-credit units, regardless of the self-recording contingency. Large increases in the percent of students moving into the medium or high groups were evident during all credit units: 50% and 45% of low-participating students in the credit units for Section A, 54% and 57% in Section B, and 27% and 54% in Section C. Although the effect of credit appears to trump the effect of self-recording, it is worth noting that the smallest increase (27%) occurred in the section in which students did not self-record (Section C). That particular unit was the first credit unit, directly after the baseline unit, which suggests the possibility of the combination of treatments (credit and self-recording together, which were operating in the first credit unit in the other two sections) having a greater immediate effect on increasing low students' participation. Still, the credit contingency appears more effective than self-recording because (a) the amount of low-participating students participating in the second credit unit in Section C approximated that of the other two sections, and (b) major changes from the low-participating group to the medium or high groups did not occur in Section B during non-credit units in which students still self-recorded participation (Units 3 and 5).

The more conservative initially low-participating group, referred to here as non-participants in Unit 1, also demonstrated increased participation in credit units, but rarely participated in non-credit units. Specifically, in Section A, 64% of non-participants commented

in at least one credit unit (50% in both credit units), while only 29% participated in a non-credit unit. In Section B, 60% participated in at least one credit unit (40% in both credit units), while 33% participated in a non-credit unit. In Section C, 57% participated in at least one credit unit (43% in both credit units), while only 14% participated in a non-credit unit. If self-recording had a significant effect on the non-participants, more of these students would have participated in non-credit units in Section B than Section C. Non-participants in Section B (who self-recorded) were slightly more likely than non-participants in Section C (no self-recording) to participate in a non-credit unit. However, because none of these students participated in both non-credit units, one might infer that the self-recording contingency was not as powerful as the credit contingency. Still, the effect of credit was not consistently strong across the non-participant group. Fourteen students, or 9% of the total sample, did not participate in any phase of the study. Stated differently, no treatment (or non-treatment) procedure evoked a single comment for one-third of the combined-sections non-participant group.

Medium- and High-Participating Students

In addition to tracking the low-participating students across units, it is helpful to see the dispersion of initially medium (mean daily participation above 0.5 and below 2) and high (mean daily participation at or above 2 comments) participants after the baseline unit. Table 13 shows this information for the medium participants (*ns* only between 9 and 11 for the three sections). In credit units, virtually all medium participants elected to maintain their participation status or to maximize the credit contingency by moving into the high group. No more than one student in each section moved from the medium to a low group during a credit contingency, with the exception of two students in Section C. The unwavering trend evident in Table 13 is that medium

participants were more likely to move into the high group in credit units, and more likely to move into the low group in non-credit units.

The majority of the initially high participants (*ns* between 17 and 23 for each section; see Table 14) stayed in the high group, irrespective of the treatment condition. Although approximately 35% or more of the high group moved into the low or medium group in at least one unit in each section, neither credit nor self-recording consistently produced this effect. For example, in Sections B and C, no more than two students moved to the low group in each unit and no more than nine students per unit (median of 5 students) moved to the medium group. However, in Section A, the combination of self-recording and credit led to slightly more students maintaining placement in the high group, in comparison to units in which both conditions were withdrawn. At first glance that result appears to run counter to what one might expect, namely that the contingencies would reduce the amount of high participants by either providing less opportunity (due to other students' increased participation) or reactivity from self-monitoring. In contrast to this expectation, these students were required to remain in the high group (comment twice per day) to earn full participation credit for that unit. Nevertheless, the inter-subject comparisons described earlier in the chapter and shown in Figure 5 demonstrated the desirable effect of credit on reducing more extreme commenting levels (5+ comments per class) that clearly exceeded credit requirements.

Type of Student Participation

Three categories of participation (comments, questions, anecdotes—previously defined in Chapter II) were recorded by the primary observer each day. The purpose of recording these dimensions of participation was to examine the possibility of students asking excessive questions

or making numerous irrelevant comments in order to earn participation credit in the course. Similar concerns were raised in previous research (Gilson, 1994; Mainkar, 2008). Mainkar used the phrase “playing for points” to refer to the practice of students getting the instructors’ attention by saying anything that comes to mind, or commenting with little regard to the contribution being made to discussion (p. 24). Although this section does not explicitly examine the quality of student participation, the type of student participation may affect the quality of the classroom discussion (e.g., excessive student anecdotes may detract from the factual nature of the information being presented). A more direct method of evaluating the quality of student responses is discussed in the next section pertaining to the assessment of student comment relevance.

The mean number of comments, questions, and anecdotes per unit in each section is displayed in Table 15. Students did not give more anecdotes or ask more questions during credit than non-credit units overall. Actually, most anecdotal responding occurred in the baseline unit. No discernible pattern exists for the number of questions asked by students under credit or self-recording contingencies. Instead, the clear trend that emerged was variable questioning and anecdotal responding depending on the content of a particular unit. Less anecdotes and more questions were asked in Units 2 (cognitive development) and 4 (psychological development) across sections of the course, suggesting that the nature of the unit content had the prevailing influence. Historically, these units have been the most difficult in the course for students and their performance has been lower in these units than in other units of the course.

Relevance of Student Comments

Assessing relevance of student comments represented an attempt to provide a direct, albeit basic, measure of the quality of student commenting. On the record cards submitted by students, there was a box next to each comment where students indicated whether each comment was relevant or non-relevant (refer back to the Method chapter for the operational definition of relevance and more procedural details for recording relevance). The primary observer also judged the relevance of student comments on her records. See Table 16 for the total number of non-relevant comments recorded by both students and the observer in each unit. Students rarely reported their comments as non relevant, and they did not consistently report fewer non-relevant comments than the observers.

No major trends in students reporting of non-relevant comments are apparent under the different treatment conditions. No claims should be made about the frequency of non-relevant claims under the different contingencies because of the small *ns* and due to the very poor agreement between students and the observer (never above 30% and as low as 0%) in the recording of relevant and non-relevant comments. Future research needs to first address the poor reliability of students' self-ratings of non-relevance—probably moreso than the definition used in the current study (inter-observer agreement was very high)—and clarify the discrepancy between student and observer perception of comment relevancy.

Interaction between Treatment and Instructor Behavior

Instructor questions. As the primary observer of instructor behavior, I recorded the number of questions (factual, comprehension, total) posed in each section on all 20 discussion days. I provided each instructor with daily feedback on their actual vs. ideal question-asking

behaviors so needed adjustments could be made to increase the equivalence between instructors and between units within sections. Instructors were taught to ask a greater proportion of comprehension than factual questions. Their behavior was largely expected to be similar due to the parallels in instructor training and the design of the course. As noted in the Method chapter, all instructors were in their mid-20s, graduate students in the same doctoral program, supervised by the same senior instructor (The supervisor met with the instructors once per week to discuss course issues and practice asking comprehension questions.) Instructors used the same course materials (e.g., syllabus, readings, study questions, exams). Also, instructors were asked to prepare discussion questions ahead of time and share them with the other instructors.

Despite all of the precautionary measures, equivalence across instructors or across units in the course could not be assumed. Thus, observing and tracking instructor behavior provided one final way to increase the likelihood of balanced instructional styles across and within the three sections through the provision of regular and individualized feedback. A consistent instructional style between units within each section was perhaps the more important issue, because most contingency comparisons were made within sections.

Table 17 shows the number of instructor questions posed in each unit for the three sections. The number of total questions and the number of comprehension questions are the primary variables of interest, the former possibly influencing the opportunities for students to participate in general and the latter possibly influencing the depth of student responses (i.e., students may need more time to formulate and deliver answers to comprehension questions and more instructor clarification may subsequently be necessary).

As an alternative to examining the sum of total questions posed by instructors in Table

17, the mean number of questions posed within each unit is displayed visually in Figure 7. Both show that Instructor B consistently posed more questions in each unit than the other instructors, and Instructor A generally posed the least. A mixed design ANOVA with section as the between variable and unit as the repeated variable revealed no significant interaction effect between section and unit, $F(8, 36) = 1.24, p = .304$. However, there was a main effect of section, $F(2, 9) = 16.06, p = .001$, and a main effect of unit, $F(1, 9) = 13.97, p = .005$. Pairwise comparisons determined that Instructor B posed significantly more questions on average per unit (38.0) than Instructor A (27.0). Significantly more questions were asked on average in Unit 1 (41.25) than in Unit 2 (29.67, $p = .005$) and Unit 5 (27.67, $p = .009$). Significantly more questions were also asked on average in Unit 4 (34.0) than Unit 5 ($p = .029$).

The cumulative number of comprehension questions posed by instructors in each unit is also shown in Table 17, and the mean number of comprehension questions in each unit is shown visually in Figure 8. The pattern is very similar to that found for total number of questions. The mixed design ANOVA with section as the between variable and unit as the repeated variable revealed no significant interaction effect between section and unit, $F(8, 36) = 1.84, p = .101$. However, there was a main effect of section, $F(2, 9) = 16.85, p = .001$, and a main effect of unit, $F(4, 36) = 7.56, p = .000$. Pairwise comparisons determined that Instructor B posed significantly more comprehension questions (25.40) than Instructor A (17.95, $p = .001$) and Instructor C (20.50, $p = .014$). Significantly more comprehension questions were asked on average in Unit 1 (24.25) than in Unit 2 (17.58, $p = .000$) and Unit 5 (18.50, $p = .010$).

Although some of the above between-section and between-unit differences were found to be significant, no pattern necessarily suggests that less or more questions were asked in any

treatment phase. The lack of interaction effect between sections and units supports this assertion. Instead, the nature of the course content in particular units may have more directly influenced the total number of questions asked. The integrity of the study would have been compromised if, for example, more questions were asked in a treatment unit than an adjacent non-treatment unit. Overall, it would be problematic if more questions were asked in treatment than non-treatment units, which hypothetically would give more students an opportunity to participate. Slightly more questions were asked in Unit 1, which was a non-treatment unit for all sections; however, this relates more to the initial workings of the feedback-to-instructors system at the start of the semester. Namely, I observed the instructor in Section C initially asking a greater proportion of factual questions than the other instructors, as well as more questions overall initially asked by the instructor in Section B compared to the other two sections. Feedback to the instructors corrected these imbalances relatively quickly. This correction is shown visually in Figures 7 and 8 as the number of total and comprehension questions asked by the three instructors converge in Unit 2, as well as numerically in Table 15 (e.g., the total number of questions asked in Unit 1 was reduced from 204 to 127 in Unit 2 for Instructor B, which was similar to the Unit 2 levels for Instructor A, 115 questions, and Instructor C, 114 questions).

The retrospective notion I developed after observing the discussion periods was that differences in the total number of questions posed by instructors were affected by the contingencies themselves, rather than participation differences and patterns being affected by instructor questioning. It was clear that students' greater participation in particular units (credit units) allowed instructors less opportunity to ask questions. Table 15 reveals that slightly more questions were asked in credit units; however, as described above with the ANOVA results, no

interaction between sections and units was significant. Because the total number of questions asked by instructors was presumed to be largely controlled by the students, the control variable that was most easily monitored and systematically controlled (equalized) across sections was the relative percentage of comprehension questions versus factual questions. Figure 9 shows that, when controlling for the slightly varied total number of questions posed, instructors asked similar percentages of comprehension questions (i.e., generally over half of the questions) within and across sections. Note that Instructor C initially asked a greater number of factual than comprehension questions, but after feedback from the observer in Unit 1 regarding the need to ask more comprehension questions, the type of questions asked by Instructor C more closely approximated that of the other two instructors. Overall, no parallels between instructor questions and treatment conditions appear to confound (inflate) the strong credit effects found in this study or minimize the weak self-recording effects.

Instructor feedback. In addition to recording instructor questions, I also recorded instructor feedback to students after each comment, coded as positive or negative. The intent was to maintain high levels of positive feedback, although this depended in part on the relevance of student comments. When non-relevant student responses were noted, negative instructor feedback theoretically should also have been recorded. However, there was slightly more negative feedback than non-relevant responses.

The proportion of negative instructor feedback to non-relevant student responses for each unit, as well as the proportion of positive feedback to relevant responses, is presented in Table 18. All of the latter proportions are close to 1.0 (ranging from .89 to 1.01), indicating that positive feedback typically followed relevant student responses. On the other hand, the average

proportion of negative feedback to non-relevant responses was more off-target (ranging from .44 to 5.00), because of the higher rate of negative feedback compared to non-relevant responses as previously mentioned. Most of these differences were not substantial, but appeared to be due to the small samples (e.g., the .44 proportion resulted from a 4/9 comparison), whereas other differences reflected considerable differences (the 5.00 proportion from a 25/5 comparison). Given the greater frequency of negative feedback than non-relevant responses it is possible this finding reflects under-reporting of non-relevant student responses, but unlikely considering the high inter-observer agreement. A more plausible explanation relates to the way in which negative feedback was defined: negative feedback was recorded when students' responses were ignored, irrespective of whether the comment was relevant or not. Thus, negative feedback appeared to generally follow non-relevant commenting, but also included occasions when no feedback followed a student response (e.g., an instructor posed a question following a student response; a student participated immediately after another student, preventing the latter student from receiving instructor feedback).

Figures 10 and 11 show the differences within and between sections regarding the average amount of positive and negative feedback given to students. Negligible differences were evident in instructors' use of positive feedback across units or sections, and only one minor difference was found in instructors' use of negative feedback. A mixed design ANOVA confirmed that differences in positive feedback between sections were not significant, $F(2,152) = 0.28$, $p = .755$, nor were differences between units, $F(4,608) = 2.30$, $p = .058$. With a small amount of overall negative feedback given in any section, instructor variation in giving negative feedback was minor. An interaction effect between unit and section was found in the mixed

design ANOVA for negative feedback, $F(8, 608) = 2.28, p = .021$. The only statistically significant difference was in Unit 3, wherein Instructor C gave more negative feedback than Instructor A ($p = .005$) and Instructor B ($p = .001$). Considering these minor differences in negative feedback and the alignment of units and treatment conditions, no definitive parallels between instructor feedback and treatment conditions exist as to their possible effects on participation. Specifically, more negative feedback was not consistently given during non-credit units (eliminating the potential of inflating the effect of credit) or during self-recording units (eliminating the potential of minimizing the effect of self-recording).

Survey Results

The first step in the analysis of the survey data was to complete a four-stage principal components analysis of responses to the 50 items in the survey. The first stage produced 14 factors, each with a minimum of 3 items loading above .30 and loading higher on that factor than on any other factor. Items that failed to meet both criteria were dropped from the subsequent stages of factor analysis. The same procedure was conducted three additional times, culminating in three factors at the end of the fourth stage. Survey item content comprising each of the factors was examined in determining appropriate factor labels: Personal History and Confidence Regarding Participation, Expectation for Discussion in College Classes, and Personal Benefits of Participation. Appendix G displays the items and factor loadings for each factor. Factor 1 is hence referred to as History/Confidence, factor 2 as Expectation, and factor 3 as Personal Benefits.

A measure of internal consistency, Cronbach's alpha, was obtained for all 50 items on the total survey (.91), the 23 items constituting the combined factors (.93), the 8 items constituting

the History/Confidence factor (.89), the 8 items constituting the Expectation factor (.86), and the 7 items constituting the Personal Benefits factor (.83). All alphas were well above the widely-accepted social sciences criterion of .70 (Garson, 2008), suggesting that each set of items could be considered a scale.

The survey data were examined in several ways: a) mean total and factor survey scores for low, medium, and high participants were compared, b) ANOVA results provided statistical backing for the differential means between the three participation groups; c) individual item means were similarly analyzed; d) content of individual items that generated extreme student responding was examined; and e) logistic regression analyses results assessed the extent to which the survey predicted placement in the low and high participant groups across units of the course.

Mean total survey and factor scores. Analysis of variance confirmed that the low, medium, and high participation groups differed significantly on the 50-item total survey, $F(2,154) = 25.3777, p = .000$. Pairwise comparisons revealed that the low, medium, and high participation means were all significantly different ($p < .05$) from each other on the 50-item total survey. Additionally, a mixed ANOVA design with the three survey factors as the repeated measure and the three participation groups as the between variable was conducted. This analysis produced a significant interaction and main effects for participation group and survey factor: interaction effect, $F(4,308) = 9.658, p = .000$, main effect for the survey factors, $F(2,308) = 17.244, p = .000$, and main effect for the participation categories as previously described, $F(2,154) = 30.724, p = .000$. Pairwise comparisons revealed significant differences ($p < .05$) between all participation groups on each of the three factors, with the exception of the medium and high groups on the Personal Benefits factor, which did not differ significantly. Similarly, the

low and medium and the low and high groups were significantly different for the discarded survey items, $F(2,157) = 7.033, p = .001$. (See Table 19 for the means of each group on the total survey, the combined factors, each factor, and the discarded survey items.)

Individual item analyses. This section examines the Participation Survey items within factors for which low, medium, and high participants scored differently. Appendix G presents the mean scores for the low-, medium-, and high-participant groups for each item on the three survey factors and discarded items. Multivariate analyses of variance permitted an examination of these item-by-item scores for the three participation groups. Specifically, one MANOVA included the 8 items on the History/Confidence factor as the dependent variables, a second analysis included the 8 items comprising the Expectation factor, and a third analysis included the 7 items comprising the Personal Benefits factor. All three multivariate analyses used the 3 participation groups (low, medium, high groups) for the total course as the independent variable. A significant overall effect for items on the History/Confidence factor, $F(16,298) = 5.927, p = .000$, the Expectation factor, $F(16,298) = 4.740, p = .000$, and the Personal Benefits factor, $F(14,300) = 3.292, p = .000$, was followed by Bonferroni post-hoc comparisons showing which item means differed significantly across all participation groups: items 1, 9, 16, and 33 from the History/Confidence factor, and item 2 on the Expectation factor.

Other item means were significantly different between the low and high groups, as well as between the low and medium groups, but not between the medium and high groups: three items (3, 4, and 38) from the History/Confidence factor; five items (8, 12, 14, 20, and 42) from the Expectation factor; and two items (22 and 25) from the Personal Benefit factor. A smaller number of item means were only significantly different between the low and high groups: item

31 from the History/Confidence factor, item 36 from the Expectation factor, and item 44 from the Personal Benefit factor. One item on the Expectation factor (34) was significantly different for the low and medium groups only. The History/Confidence and Expectation factors both yielded some significant group differences for each item. Four items (21, 23, 24, and 29) on the Personal Benefit factor failed to predict any differences between the three groups.

Irrespective of participation group, examination of the overall Participation Survey item means, the percent of students who chose each option, and the content of these items (Appendix E) revealed additional information regarding student perceptions of participation. As described in the Method chapter, item responses were coded between 1 and 5, with higher scores indicating a more favorable perspective of participation. Almost half of item means were mid-level (between 2.5 and 3.5), 32% of item means were high-level (between 3.5 and 4.0), and 22% of item means were very high (above 4.0). Interestingly, no items that make up the latter group constitute any part of the three factors. It may be that these items do not distinguish the low and high participant groups as well as some of the others.

Results of the items with very-high means suggest the following patterns: students perceive their own comments to be always or generally relevant (item 5); students feel responsible for participation when instructors ask students to volunteer (item 7); students report positive feelings about participation being welcomed but optional (item 10); students prefer volunteering comments rather than being called on (item 17); students expect peers who participate frequently to do well (item 27); students comment when they feel knowledgeable on a topic (item 28); students recognize discussion is valued by teachers (item 37); students believe they can judge the relevancy of their own comments (item 41); students desire to participate in

courses they view as having relevant content (item 47); students perceive that participation results in greater liking by teachers (item 48); and students regard teacher friendliness as an important contributor to student participation (item 49).

No items had low-level means. The four lowest item means were below the neutral score of 3.0 but still considered mid-level, and the results of these items suggest the following: approximately a quarter of students think participation should affect grades minimally (item 8); students prefer a balance between lecture and discussion, with a slight leaning toward more lecture (item 14); students have mixed feelings about the relationship between quantity and quality of participation, but slightly more students regard quantity as detracting from quality (item 15); and students do not admire quiet students but rather regard them no differently than other peers or may tend to discredit them (item 46).

Logistic regression analyses. The predictive potential of the three factors, as well as the total survey, was assessed by completing a series of logistic regression analyses. The intent was to see how well the instrument could predict placement in low versus high participation groups for each unit rather than only for the composite units. The particular students in the two groups varied from unit to unit: low participants (0-2 comments for the unit) and high (8+ comments for the unit). The cutoff criteria represented a substantial separation between the low and high groups in each unit: the low group commented no more than twice during each four-day unit and the high group averaged commenting at least twice each day.

The classification levels used in this set of analyses are different than those used to evaluate daily changes in participation; however the classification is consistent with how the low- and high-participating students were identified in Unit 1 for the intra-subject comparisons.

The rationale for excluding the medium cases (i.e., students making 3-7 comments for the unit) from the logistic regression analysis was twofold: (1) the *ns* for the medium participants were generally smaller than for the low and high groups, especially in non-credit units, which would result in the in the logistic regression analyses' allocating fewer cases to the medium group, blurring differences between the low and high groups, and (2) constructing low and high groups permitted the exclusion of the most variable participants for whom the total survey or any survey factors would likely not predict participation well.

Binomial logistic regression allows evaluation of the contribution made by each predictor (survey factor) over and above that of the other predictors. First, the total survey scores (based on all 50 survey items) were evaluated as predictors of low versus high participation classification. Next, logistic regression was done for the three survey factors combined. The model evaluated the combined factors as a predictor before examining how each factor separately contributed to the model. Finally, the factor that was clearly superior to the other two factors in predicting high and low participation was compared with the total-survey model and the combined-survey model in predicting membership in the high and low participation groups.

The logistic regression analyses generated several statistics. Chi-square values and corresponding *p* probabilities, as well as Cox and Snell R^2 , assessed the strength of the overall model (i.e., the total survey versus the 3 factors versus individual factors as predictors of participation). Note that although the R^2 available in logistic regression is a possible measure of effect size, that R^2 is considered a pseudo R^2 and should be interpreted with caution (Garson, 2009). Wald statistics and corresponding *p* values, *B* coefficients, and odds ratios permitted examination of the predictive potential of the individual survey factors. Wald statistics are chi-

square values that test the significance of the individual factors as contributing to the overall model. B coefficients are values for the logistic regression equation for predicting the dependent variable from the independent variable. The B coefficients are the natural logs of the odds ratio, in log-odd units, and are directly related to and more easily interpreted by the odds ratio (i.e., when $B = 0$, the odds ratio = 1, indicating no predictive relationship). Specifically, the odds ratio reveals how much more likely a student would be to be classified as a high participant with every 1 unit increase in the survey score. Each of these statistics are presented in Tables for each unit, as well as for the various combinations of survey items (i.e., the total 50-item survey, items from the three survey factors, and items from the History/Confidence factor).

As shown in Table 20, total survey scores significantly predicted student classification into low or high participation categories in all five units. Chi-square values were statistically significant ($p < .001$) in each unit, with R^2 values ranging from .18 to .34. Odds ratios ranged from 1.05 to 1.10, indicating that students were 5-10% more likely to participate (i.e., be included in the high participation group) with every one unit increase in the total participation survey scores. Lower than expected odds ratio values resulted from the high degree of variance associated with the high participation group. Specifically, the range of participation per unit in the low group was only 2 (0 to 2 comments for any student), whereas the range in the high group was 36 (no lower than 8 comments per unit for any student, but as high as 44 comments). The highest R^2 and odds ratio values were obtained in Units 1 and 2, suggesting that the 50-item survey best predicted student participation levels near the beginning of the course. Perhaps in subsequent units other factors (e.g., peer relationships, comfort in participating, perception of the

need to comment to boost a grade or no need to comment due to solid grades) carried added weight.

Logistic regression allowed for the examination of group differences not only statistically, but practically as well. The analysis classified cases (students) into either the low or high participation group based on their total survey scores. The percent of students correctly classified for all sections combined was between 67.2% (Unit 3) and 77.6% (Unit 2), indicating a high degree of accuracy in predicting membership in the low and high groups overall. (See Table 21 for the percent of cases correctly classified, with percentages provided for both low and high groups across units.) In the main, the total survey scores predicted membership in the low-participation group (percentages ranging from 63.6% to 88.9%, with a 74.2% mean across units within sections) and the high-participation group (percentages ranging from 43.8% to 85.7% , with a 69.5% mean across units within sections) equally well ($p = ns$).

Logistic regression analyses were also completed for each unit using only the survey factors as predictor variables. The combination of factors significantly predicted membership in the low and high participation groups in all five units ($p < .001$, R^2 ranging from .18 to .45). Again, the highest effect sizes were found in Unit 1 (.45) and Unit 2 (.38). To assess the relative importance of the three survey factors, the odds ratios and classification results are presented (see Table 22 for all relevant statistics). The R^2 values for the combined-factor model (Table 22) appeared higher than the total-survey R^2 for 4 of 5 units (Table 20). The History/Confidence factor was the only factor that contributed significantly ($p < .05$) to the predictive potential of the three-factor model in distinguishing between low and high participants across all five units. The odds ratio values corresponding to the History/Confidence factor were between 1.13 and 1.32

across units, meaning that students were 13-32% more likely to be included in the high participation group with every one unit increase in the combined factors survey scores. The Expectation factor was a significant ($p < .05$) predictor variable only in Unit 1, with an odds ratio of 1.24. The Personal Benefits factor was not a significant predictor variable in any unit. In fact, the B coefficients were very close to zero (some were negative values) and the odds ratios were close to 1, suggesting that the Personal Benefits factor did not contribute to placement in the high and low participation groups.

Additionally, Table 23 displays the classification percentages for the three-factor model. Its percentage of correct placement in the low-participation groups was exactly the same (74.2%) as had been obtained for the total survey. However, the percentage of correct placement in the high-participation group was higher (79.0%) than was obtained for the total survey (69.5%). Overall, the percentage of correct placements in the two participation groups was 76.7% compared to 72.3% for the total survey.

The final series of logistic regression analyses included only the History/Confidence factor, which was the exclusive or primary significant predictor in all units. Table 24 displays the statistics for the History/Confidence predictor in each unit. Very minimal differences are evident in these results compared to the analyses with all three factors combined. Table 25 presents the classification results for only the History/Confidence factor model across units. This factor produced essentially the same accuracy of prediction for the low-participation group (73.5%) as was the case for the total-survey and three-factor models (74.2%). The History/Confidence factor manifested much the same level of correct placement in the high-participation group (77.9%) as was achieved in the three-factor model (78.9%) and the total-survey model (72.3%).

One caveat in interpreting the logistic regression results is that when predictor variables are highly correlated, the outcome may show that a predictor that has a strong correlation with the dependent variable has no predictive value in the presence of the other predictor variables. In fact, correlations between the three predictor variables were all significant ($p < .01$): .55 for History/Confidence and Personal Benefits, .69 for Personal Benefits and Expectation, and .75 for History/Confidence and Expectation. Nevertheless, the Personal Benefit factor was less strongly correlated with participation level (.35) than were the other two factors: Expectation and participation level (.59) and History/Confidence and participation level (.66). In sum, the inclusion of the Personal Benefit factor items would likely not diminish the predictive potential of the survey but would add little to the prediction of low and high participation. The Expectation factor slightly increased the predictive capacity of the survey early in the course (i.e., the first four days in this study).

Because items on the History/Confidence factor had the highest predictive potential of the three survey factors and predicted participation level as well as the three survey factors combined, we closely examined its potential for predicting low/high participation placement when credit was given for participation compared to when credit was not given for participation. My hypothesis was that the self-report participation survey would predict participation better when no credit was given for participation than when credit was given. This hypothesis should be especially applicable to the strongest predictor, the History/Confidence factor. Although credit for participation could be expected to account for some of the variation in participation rates during credit units, the History/Confidence factor should account for more of the variation in participation during non-credit units in which no added incentive was applied to participation.

Table 25 identifies the credit and non-credit units for each section of the course. The results did not confirm overall better predictive potential for the History/Confidence factor in non-credit than credit units, but rather better predicted placement for the low-participant group than the high group in the non-credit units, and better predicted the high-participant group than the low group in the credit units. As previously described, there were more low participants during non-credit phases of the study. Logistic regression analysis showed that the History/Confidence factor correctly predicted this trend: more students were predicted to be low participants in the two non-credit units across sections (151) than in the two credit units (76). Yet, in addition to predicting more low participants in non-credit units, the factor more accurately predicted low participants in non-credit units (81%) than in credit units (71%). Use of Ferguson and Takane's (1989) proportion's test showed the differences in these proportions to be statistically significant ($p < .05$). The opposite trend occurred for high participants. The History/Confidence factor better predicted placement of these students in the credit units (89%) than non-credit units (65%). A test of proportions determined these differences were statistically significant ($p < .0001$). In some non-credit units, the percent of cases classified correctly into the high group was no greater than what would be expected by chance (e.g., only 52.9% in Unit 3 of Section C). On the other hand, for many of these credit units, the percent of cases classified correctly was notably higher for the high group than the low group (e.g., 91.7% in Unit 4 of Section B, and most other percentages approaching or exceeding 90%).

Chapter IV

Discussion & Conclusions

This study evaluated the efficacy of crediting participation and requiring students to self-record their daily participation in the quest to balance participation among students in three sections of an undergraduate course, each with approximately 55 students. In a large discussion-based course with considerable disparity among students with respect to their initial level of class participation, the principal goal was to better balance participation across students without diminishing the relevance of discussion rather than to merely increase the quantity of participation. Instructors of lecture-taught classes are more likely to focus on increasing quantity of participation; the quantity dimension may not be as relevant in discussion-based classes because there is a ceiling for participation that relates to the number of instructor questions asked and length of the class period. In addition to determining the effect of credit and self-recording on participation, a survey was developed to identify characteristics of students participating at different levels.

Results revealed that the crediting procedures were largely effective in balancing participation. The percent of participants in each of the three sections of the course was significantly higher in credit than non-credit phases (recall that each phase was referred to as a unit in the course, comprised of four discussion days in each of the five units). Specifically, the percent of non-participants was reduced approximately 50% in credit units; the percent of frequent participants was relatively similar in credit and non-credit units; and the percent of dominant participants was reduced from approximately 9% in non-credit units to 3% in credit

units. More detail on the effect of credit on students' participation is specified below according to the levels (low, medium, high) at which they participated.

Effect of Credit on Patterns of Student Participation

To determine the effects of participation credit, students were classified by (1) level of participation per day within each unit and (2) range of comments for the unit as a whole. Four categories for percentage of students participating each day were designated: non-participants (0 comment each day), occasional participants (1-2 comments each day), frequent participants (3-4 comments each day), and dominating participants (5+ comments each day). As previously noted, students were independently selected for the different levels each day without regard to their level of participation on other days. This classification scheme permitted examination of patterns of daily student participation. In contrast, the classification of students by unit resulted in three different levels: low participants (0-2 comments for the unit), medium (3-7 comments for the unit), and high (8+ comments for the unit). These criteria allowed for more global comparisons to be made, including predictability of the survey data between treatment and non-treatment units (i.e., the tracking of the participant groups from unit to unit).

The most consistent pattern discerned from the daily participation categories was for the non-participant group. The daily percentage of non-participants varied minimally within units (credit or non-credit) despite the large variability in percentage of non-participants between credit and non-credit units. On the other hand, the percent of frequent participants (3-4 comments per day) was variable across days within and between units. The findings for the dominating participants were as clear and compelling as for the non-participants. There were less than five percent of dominating participants on each day in credit units, and much higher percentages on

days in non-credit units. Specifically, the average percent of dominating participants across non-credit units was 9% and was as high as 16% on one day.

Effect of Credit on Low-Participating Students

Henning (2005) reminded instructors of the need for direct strategies to facilitate participation from highly reticent students: “No matter how well framed or well facilitated, the discussion still may not work if students are too reticent or reluctant to share their thoughts publicly. When this is the problem, making the questions even more thoughtful will not help; more direct strategies are needed” (p. 93). Thus, a principal goal of this study was to offer a direct strategy that would move non- and low-participants to the medium- and high-participation categories. Direct strategies used in the current study were credit for participation and self-recording of participation. Without either of these strategies, approximately half of students did not participate at all during the initial unit of the course; under credit conditions slightly more than 90% of the students participated in one or more credit units.

Not all studies define reticence similarly. It may be that reticence is equated with non-participation, little to no participation on average, or little to no participation per class period. Justification for the latter comes from studies evaluating the consolidation of responsibility norm (Howard et al., 2002; Karp & Yoels, 1976). Howard et al. suggested that the operation of this norm, referring to only a handful of students accounting for the vast majority of interactions in any given class session, produces only “talkers” and “nontalkers” in the classroom. Talkers were defined as students who account for the vast majority of all interactions (operationally defined in their study as 2 comments and above per class) and nontalkers were defined as students who speak only occasionally, if at all (contributing less than 2 comments per class in classes of

approximately 20 students). Thus, they argue that there are no “average” participants; talkers and nontalkers should be differentiated each class session, because looking at mean statistics blurs distinctions between students who accept and reject the consolidation of responsibility.

Although dividing levels of participation into two groups (talkers and nontalkers) as done in Howard et al. (2002) has some benefits, I argue that it is reasonable to differentiate three groups of participants (low, medium, and high) in certain circumstances. The classrooms in the current study had slightly over 50 students, the instructional style (discussion-based with little to no lecture) required a consistent and high class-wide rate of participation, and a direct strategy (attaching credit to participation) was applied to balance participation among students. The goal of the intervention, balancing participation, would eliminate or reduce the consolidation of responsibility. Approximately half of students or more typically participated at least once in each session and, for example, on one day as few as 8-10 of the 50+ students in each of the three sections of the course failed to participate. Most students participated on credit days and at about the same level, inasmuch as all students were expected to make 2 comments to earn full credit. Using the criteria in Howard et al. to define a talker group in the current study would have resulted in an excessively large nontalker group. However, reducing my criterion for the talker group to 1 comment per day would fail to differentiate between students who spoke occasionally from students meeting and exceeding the credit requirements. Instead of the talker-nontalker categories, it was more meaningful to differentiate students who rarely or never participated (0-2 comments per unit), from medium-participating students (3-7 comments per unit), and high-participating students (8+ comments per unit, or an average of 2 or more comments each day

within a four-day unit). In this study, reticence is defined as the first group, which adheres to the guideline proposed by Howard et al. for students who speak up only occasionally, if at all.

Reticence was defined in multiple ways in the current study: initial reticence, defined by low participation (0-2 comments) in Unit 1; persistent reticence, defined as low participation in at least four of the five units; daily reticence, defined as the percent of non-participants each day; and history of reticence as defined by grouping students who scored low on the History/Confidence factor items on the Participation Survey. Regardless of which way the data were analyzed, the crediting procedure was effective in substantially reducing student reticence, although some (9% of the total sample) remained non-participants throughout the course.

The percent of initially reticent participants ($n = 67$), participated more in credit units (46%) than non-credit units (13%). On average, about half of the initially low-participating students moved into the medium- or high-participant groups during credit units, while the overwhelming majority remained in the low group during non-credit units. Notably, the subset of this group who did not make a single comment during the first unit ($n = 43$) also had a higher percentage of participants during credit units (between 40 and 50% participated in both credit units) than in non-credit units (between 14 and 33% participated in a non-credit unit). Inter-subject comparisons revealed that the percentage of non-participants was cut nearly in half in credit units, and there were significant differences in the proportion of non-participants in all sections of the course from each credit to non-credit unit.

It is reasonable to conclude that the crediting procedures were effective in increasing the level of participation of initially low- and non-participants, as well as reducing the percent of non-participants overall. What adds weight to this contention is the regression of these students

to low levels of participation in subsequent non-credit units. Moreover, of the students who scored in the bottom third on the survey items from the History/Comfort factor (the strongest factor), approximately twice as many participated in credit units as non-credit units. However, 23% of all students remained low participants in at least four of the course units, and 9% of the entire sample never participated in the course. It is the latter students who most need to be carefully tracked in the classroom and their perception and history of participation better understood.

Effect of Credit on Medium-Participating Students

The medium-level participants ($n = 31$) in the first unit typically participated at the medium or high level in subsequent credit units but at the medium or low level in non-credit units. A contrast effect between credit and non-credit units appeared to make participation less desirable in non-credit units for initially medium participants. Also, in Sections A and B, initially medium participants were more likely to remain at the medium level than to become high participants in credit units. In contrast, in Section C, slightly more of the initially medium participants became high participants in credit units than remained medium participants.

The effect of credit on medium participants is somewhat of an anomaly. Some students in the medium group participated consistently across days but at the lower level of the credit contingency (e.g., one comment per class, compared to the contingency requirement of two comments to maximize credit). For other medium participants, credit produced variable participation across days. In either situation, the result was continued classification as medium participants. As described earlier, Howard et al. (2002) proposed that no “average” participant

exists. Thus, medium participants should not necessarily be likened to average or regular participants, but rather variable participants.

Effect of Credit on High- and Dominating-Participants

Most high participants in Unit 1 participated at a high level in the subsequent units, though some did become medium participants and a few even low participants. Because two comments a day were required to maximize participation credit, high participants had to stay at this level to earn maximum participation credit. Nonetheless, the high-participation percentages for this group differed minimally between credit and non-credit phases. The high participants appeared to be the one participation group for whom sources other than the credit sustained their participation, perhaps either the intrinsic satisfaction of participating, the positive effects of participation on learning, or the social reinforcement for participating.

In this study, high participants were differentiated from dominant participants. As previously noted, a student could be classified as a high participant by meeting the minimum participation requirements to receive full course credit (i.e., 2 comments per class). On the other hand, dominating participants commented five or more times per class on average. Not surprisingly, credit for participation greatly reduced the percent of dominating students in credit units. Interestingly, this trend was more apparent in the two course sections that required self-recording during the credit units. In Sections A and B, there was never more than one dominating student in any unit involving both credit and self-recording. Credit reduced the number of dominating students from non-credit to credit units (from about 6 to 3 on average) in Section C; however, significant differences in proportions of dominating students from credit to non-credit

units were found in only the former two sections. Perhaps self-recording had a reactive effect on extremely high levels of commenting.

Reliability and Effect of Student Self-Recording

Unlike the effect of the crediting procedure, the self-recording procedure had a negligible effect on balancing participation across students. Still, a possible effect of reducing the percent of dominating participants deserves consideration. Additionally, and perhaps more importantly, self-recording proved to be a reliable and efficient means of quantifying student participation in order to assign course credit. This latter finding may be surprising for instructors who are familiar with previous research. Correlations between instructor and peer assessments were between .83 and .90 in one study (Melvin, 1988), which Mainkar (2008) notes is much higher than correlations between instructor and student self-assessments as found in a meta-analysis conducted by Falchikov and Boud (1989). Also, students have been noted to over-report their participation during end-of-class assessments (Burchfield & Sappington, 2000; Gopinath, 1999). On the contrary, this study shows that students are capable of accurately reporting their daily participation. In fact, students were more likely to under-report their participation compared to observer records. Correlations between student and instructor daily records were very high, ranging from .86 to .99.

Similar levels of student-observer agreement were found in Krohn et al. (2010) when students used the same record card system as used in the current study. Despite the high levels of agreement in the Krohn et al. study and the current study, students tended to over-report their participation in credit units and under-report their participation in non-credit units. Instructors desiring to use the self-recording system to credit students for participation throughout a class

should be aware that some students may abuse the system to receive undue credit. However, this type of over-reporting rarely occurred in the current study. Although significantly more students over-reported their participation in credit than non-credit units, no more than 2 to 7% of students per day engaged in this practice. Furthermore, only five total students over-reported to receive credit on more than one of the eight credit days; the other sixteen students who received undue credit only over-reported on one occasion. It is unlikely that over-reporting for credit was intentional, and it was not excessive for any student. The demographic data of the five students who received undue credit on more than one day revealed no discernable patterns: 3 students were male and 2 were female, GPAs ranged from 2.4 to 4.0, pre-course scores ranged from 26% to 60% (mean score for the total sample was approximately 44%), and their course exam average ranged from 75% to 93%.

Students under-reported their participation significantly more in non-credit than credit units. This tendency rarely impacted the amount of credit students received for participation because students generally only under-reported above the criterion for full daily credit (i.e., after recording 2 comments). Instructors should be aware that students may under-report their participation if no credit is provided. Yet, in the main, this study lends support for using students' daily self-recording as a basis for giving credit for participation. Krohn et al. (2010) did not find significant differences between the patterns of over- and under-reporting in credit and non-credit units. Although significant differences were found in the current study, under- and over-reporting occurred at low rates overall. Students did not appear to be intentionally abusing the system to receive undue course credit.

Another possible undesirable side effect resulting from students under- or over-reporting, even if not extensive, is reduced internal validity in the study. For example, students' over-reporting in credit units and under-reporting in non-credit units could inflate the treatment effect of credit. Fortunately, this problem was bypassed in the current study because observer records were used as the main dependent variable throughout. However, those desiring to conduct research on crediting and/or self-recording procedures should be careful to avoid overstating the likelihood of a treatment effect if no observers simultaneously collect the data.

As noted previously, self-recording did not appear to affect levels of participation directly. Changes in percent of participants across units in the three sections were consistently better explained by credit and non-credit unit differences rather than differences in whether students self recorded their participation. One exception was a lower percent of dominating students (5+ comments on average per day) in credit units with self recording than credit units without self-recording. There is some evidence that self-recording would be beneficial for dominating participants: the only section (Section C) for which credit did not affect the percentage of dominating students was the section in which students never self-recorded their participation. In the main, fewer dominating students participated when both credit and self-recording were used in the other two sections; however, this treatment effect was not consistent (i.e., no significant difference was found in the percentage of dominating students from the last credit unit to the last non-credit unit). Thus, caution should be exercised in assuming that credit for participation will reduce the percentage of dominating students.

Non-effect of Differential Instructor Behavior

Just as students' under- and over-reporting of participation needed to be minimized as tendencies that could inflate the treatment effect of credit, differential instructor behavior could also have an extraneous influence on student participation. Consequently, both instructor questioning and feedback to student comments were monitored. No significant interaction effects between units and instructors were obtained for either type of questions or type of feedback (i.e., instructor behavior did not differ significantly under different treatment conditions). Treatment effects of credit could have been inflated if, for example, more questions were asked in credit than non-credit units, which presumably would allow for more opportunities for students to participate. Instead of instructor questions inflating the treatment effect of credit, the opposite may have occurred: instructor questions may have diminished during credit conditions because students were participating more. Stated differently, it may be that crediting participation resulted in more active participation by students, pre-empting some time for instructor comments.

In addition to carefully orchestrating their questions to facilitate discussion, instructors should also be aware of the potential effect of feedback on student participation. Operant conditioning principles suggest that differential types and amount of feedback alone could impact student participation with or without formal crediting procedures. In fact, Hodge and Nelson (1991) reported that using differential reinforcement in the form of marks on the blackboard next to student names increased or decreased student participation in the expected direction, ultimately creating a better balance in participation across students. The Hodge and Nelson study underscores the power of differential feedback in increasing or decreasing

participation. Instructor feedback appeared not to be a confounding variable in the current study, largely because instructors were trained to provide positive feedback to students across circumstances, whatever experimental condition was in effect. That being the case, a very high frequency of positive feedback and a very low frequency of negative feedback occurred across sections and units within sections. No significant interactions occurred between sections and units in the use of positive feedback. Some differences were found in instructors' use of negative feedback but not to the extent of enhancing or diminishing the power of any treatment condition.

Effect of Credit on Relevance of Participation

As previously described, the effect of credit on balancing student participation levels was consistent and strong across the course sections. The findings presented for quantity of participation as the dependent variable are relatively unequivocal. However, an ancillary goal was to determine the collateral effects on the relevance of student comments. Although no credit contingencies were attached to the relevance of participation, changes in levels of participation could indirectly produce qualitative changes. Balancing quantity of participation across students might also improve content relevance. On the other hand, increasing the percentage of students commenting in the discussion could also weaken relevance of students' comments as they rushed to earn credit for participation.

The first step in examining the effect of treatment conditions on quality of participation was to operationalize and assess quality of participation. The assessment of quality yielded poor reliability between student and observer records, although not between observer records. Students and observers agreed on relevance ratings to a considerable degree (correlations between .77 and .99). However, students and observers rarely agreed on non-relevant ratings.

Despite high inter-observer agreement on non-relevance of comments, agreement between students and observers within sections and units ranged from 0% to 30%. Few instances of non-relevant comments occurred according to observer and student records. For example, the lowest agreement level in a unit was 0 out of 16 times, and the highest agreement level in a unit was 8 out of 27 cases. Observers typically recorded the higher number of non-relevant cases.

One suggestion for enhancing the study of comment relevance or quality would involve stronger emphasis on accurately self-recording quality of participation, as well as discussing with students how and why high-quality student participation will benefit their course experience. As a side note, this suggestion also pertains to the introduction of self-recording of quantity. Students in the current study understood that information on the record cards would be the basis of course credit in self-record units. However, they were not given a detailed rationale for why the record card system was used. If students understand that the record card system will allow the instructor to objectively credit their participation and to better understand the dynamics in the classroom and the extent of balanced participation at any time, they may be more receptive to self-recording. After laying a foundation for the importance of accurate quantity ratings, instructors can explain that they are also interested in students' ability to self-assess the quality of their participation and to strive for consistently high self-ratings. Increasing student investment in these procedures would likely increase both the fidelity and effectiveness of self-recording.

I cannot make any solid claims regarding the effect of credit on quality of participation due to poor student-observer agreement levels for the relevant dimension and the low overall frequency of non-relevant ratings. Nevertheless, the low frequency of non-relevant commenting

overall attests to the unlikelihood of credit producing poor-quality participation; because quality of participation can be defined in many different ways, a more cautious statement would be that crediting participation is not likely to produce more non-relevant commenting. Observers recorded low rates of non-relevant responding throughout all phases of the study, discounting the undesirable possibility that crediting participation would result in more inaccurate, redundant, or off-topic comments. Additionally, the effect of credit on a related but indirect measure of quality was also determined: type of student participation (comments, questions, anecdotes) as recorded by observers. Students were not more likely to ask excessive questions or give more anecdotes during credit phases, which could also diminish the quality and effectiveness of the discussion.

Given that non-relevant commenting was rarely reported by both students and observers, quality of participation as measured in this study may have been too narrowly defined. Perhaps students are capable of making finer distinctions as to whether specific comments contribute to the class discussion. If students are asked to make dichotomous judgments (e.g., relevant vs. non-relevant), self-serving biases might cause students to rate themselves favorably. After all, why would students knowingly offer inaccurate, redundant, or off-topic comments that would be poorly received by the instructor? It may be more meaningful to students to rate the quality of their comments on a graded scale, rather than the narrow relevant vs. non-relevant dimension. Although grading ratings might be more difficult, such a scale would probably give students greater insight into the contributions they are making to class discussions.

Mainkar (2008) suggested three categories be used to judge the quality of participation: no-substance (e.g., does not add to the understanding of the topic—redundant, superficial, irrelevant), straightforward (e.g., adds to the understanding of the topic, provides enough content

to answer the question with or without some superficial or irrelevant content), and insightful comments (e.g., significantly improves the understanding of the topic, shows substantial depth, makes creative connections). The distinction between the latter two categories may prove especially helpful—a three-category quality dimension may improve the reliability between student and observer records.

My prediction is that reliability would be much higher with three or more categories than the two used in the current study because students would be less apt to default to the positive and socially acceptable (relevant) rating. Instead, more effort may be put into consideration of whether they personally responded insightfully or straightforwardly, which is a more cognitively-demanding task than parsing out relevant from non-relevant comments. I propose that an even wider scale be used, with possibly four or five brief descriptive categories, to further expand the effort with which students categorize their comments and produce more variability in ratings. Greater effort in categorizing their comments may increase accuracy and contribute to comments that meet the highest quality-rating criteria. At the very least, Falchikov and Boud (1989) claimed that self assessment can be a valuable learning activity, even in the absence of agreement between students and instructors (or observers).

First and foremost, directly addressing and practicing self-assessing quality of participation is needed at the beginning of a course. Self-assessment should be regarded as a skill that, like any other skill, needs to be developed (Falchikov & Boud, 1989). Other things must also be considered when studying quality of participation, such as accounting for difficulty of instructor questions. Extraneous factors can be more easily controlled within a study than between studies. The latter constitutes an additional source of error that may occur in integrating

others' findings on the effect of self-recording quality of participation. Additional considerations include instructor feedback, definition of the quality construct, explanation of that construct to students, and opportunities for students to practice and receive feedback on their self-ratings.

A direct connection between grading participation and discussion quality was made in at least one study (Dallimore et al., 2004), in which students indicated that they believed grading participation would enhance the quality and effectiveness of discussion. It should be noted that the study was conducted in graduate business classes, in which a high rate of participation was expected. Future research should aim to untangle the seemingly contradictory findings in the assessment of student perceptions of participation. The current findings cast doubt on the prospect of improving quality of participation as a by-product of increasing quantity. Quality of participation may be unlikely to increase unless credit is given directly for quality. If a measure of quality or relevance can be developed with sufficient precision that students can reliably rate quality, then credit contingencies can be attached to quality of participation. Nonetheless, the development of a measure of quality that can be used reliably by both students and observers is a tremendous challenge and is the necessary first step for other research questions surrounding the quantity-quality link.

Predictors and Perceptions of Participation

A major aim of this study is to provide instructors with a practical, effective and efficient way to enhance classroom participation for sizeable groups of students. Although providing credit for participation will increase the percentage of students participating in discussion, some students will likely be unresponsive to that credit. In medium or large classes, discerning individual differences among students may help account for their level of responsiveness to

participation credit. Thus, a secondary purpose of the current study was to develop a participation survey that would illuminate individual differences affecting participation. This information may help instructors develop instructional procedures that will encourage the most reticent students to become involved in discussion.

Logistic regression results showed that the Participation Survey was a strong predictor of which students would manifest low and high participation rates. Three factors emerged from the Principal Components analysis of the survey items: Personal History and Confidence regarding Participation (History/Confidence), Expectation for Discussion in College Classes (Expectation), and Personal Benefits of Participation (Personal Benefits). The combination of these three factors produced an R^2 generally equivalent to that of the survey as a whole. Of the three factors, the History/Confidence factor was the strongest predictor of participation. In fact, the History/Confidence factor predicted participation as well as the total survey and the three factors combined. The Expectation factor only added to the predictive potential of the History/Confidence factor in Unit 1, and otherwise did not add to the predictive potential of the survey. The items on the Personal Benefits factor did not add to the predictive potential of the survey. Thus, instructors looking for a parsimonious, efficient survey to predict student participation may choose to administer only the eight items from the History/Confidence factor. However, administering the items on the other factors would not lower the predictive capacity of the survey and might provide additional insight as to what could motivate students to participate.

In terms of practical usage of participation surveys, if an instructor's primary concern is predicting who will be the low participants, the History/Confidence factor may do better under non-credit than credit conditions. Thus, administering the brief 8-item survey may be particularly

useful to instructors disinclined to provide formal credit for participation. On the other hand, if the instructor's primary interest is predicting who will be high participants, then the History/Confidence factor will predict much better under credit than non-credit conditions.

History/Confidence indicators of student tendencies to participate are more likely to be activated by credit than non-credit conditions, whereas History/Confidence indicators of disinclination to participate are more likely to be influential under non-credit than credit conditions. The best explanation for this result may be that credit was effective in propelling some low participants into the medium or high groups. Stated differently, the effect of the credit explains why History/Confidence is a better predictor for the low group in non-credit units. In credit conditions there was a prominent other variable (credit) accounting for change in participation for some low students. The movement of some low participants to medium or high groups blurred differences (e.g., in history, comfort, confidence, perceived value of participation) between the groups that existed in a baseline/non-treatment condition. Thus, low participants may have entered the course with varied contributions to their history of non-participation, for which some were more easily reconciled than others with the small incentive of credit. The low participants under the credit condition were likely more hard-core low participants than the low participants under non-credit conditions. A student with a history of low participation and confidence in participation who experiences slight discomfort when speaking in public, for example, could be separated from a student with a similar history emanating from more substantial issues (e.g., major anxiety, frequent experiences of poorly received previous attempts to participate).

High participants were much more easily identified by the History/Confidence factor in

credit than non-credit units. High participants may have an inclination to participate, presumably due to a longstanding history of participating and confidence in doing so. Credit appears to accentuate that tendency, likely by establishing the norm that it is okay to participate. Credit for participation communicates the expectation of participation, whereas social norms in the classroom may promote the opposite under normal/non-credit circumstances. Smith (1992) suggested that students who oppose credit complained that it seemed too much like cooperation with the instructor for the sake of a grade. He further stated that the norm appeared to be for students in his large classes to do only what was required and not one bit more. Thus, in the current study, the presence of credit for 2 comments per class set an expectation for participation that would prevent students from being perceived as obsequious. Even students with a strong history of and confidence in participation may limit their contributions based on classroom norms. In non-credit conditions, high participants' history/confidence is muted by the desire to follow the social norm. It is no longer expected or desirable to participate at high levels, leading to more variable participation for this group. Another explanation for the factor's better predictability in credit than non-credit units may be that History/Confidence has a correlate: personality-related characteristics, such as high achievement motivation, which produces consistently high participation to maximize the number of points earned in the course.

Possible sequencing effects may have contributed to an overjustification effect, serving as a confounding variable or simply accentuating the better predictability of high participants in credit conditions. That is, credit conditions and non-credit conditions were always alternated in the current study. A possible explanation for better predictability of high participants in credit conditions may relate to setting the expectation for everyone to participate and subsequently

removing that expectation. In non-credit units (baseline excluded) the History/Confidence factor predicted high participants only slightly better than chance would predict. Credit may or may not have influenced students with strong histories of participation to initially contribute to discussion, but removal of the credit contingency may have communicated to these students that regular participation was no longer expected. Even more likely is that removal of credit communicated the message that regular participation was no longer required and students would then perceive regular participation as going above and beyond the instructor's expectations.

The combined results from the logistic regression analyses suggest that the best predictor of student participation in a course (comparable to the one in the current study) may be previous patterns of participation and comfort or confidence when contributing to class discussion. Unfortunately, the opportunity that an instructor has to change patterns of participation in a course is restricted. History of participation cannot be reversed. Thus, instructors should focus on increasing students' confidence in contributing to class discussion. Instructors should consider adjusting their approach to teaching to facilitate relationship development with students (e.g., having informal discussions with students before and after class, learning their names) and provide students opportunities to increase their familiarity with the material before class (e.g., required readings and study questions, prompts at the beginning of class with which students discuss in small groups before the class convenes as a whole). However, another approach to increasing students' confidence is to try to increase the frequency with which they participate. Participation can be considered a skill, which requires practice to develop and which eventually produces confidence in participation.

Examining responses to particular survey items appeared especially relevant in further isolating important predictor variables, as well as better understanding and accounting for participation outcomes in the current study. First, according to the survey, relevancy distinctions should have been difficult for only 3.1% of students; however, the agreement between student and observer ratings of relevance was very low and even nonexistent in some cases. Secondly, one item highly endorsed by students regarded the relationship between perceived relevance of course content and student inclination to participate: students feel the greatest desire to participate in courses they view as highly relevant. Unfortunately, students may enroll in large or introductory college courses because they are required and not because they reputedly have interesting content. Despite instructor effort to individualize and stimulate discussion, not all students will find the content relevant. For these students, credit contingencies or related participation incentives may be especially important in promoting participation. Not surprisingly, students more likely to find the content relevant (i.e., students accepted to the University's teacher preparation program) participated at higher levels (approximately 75% in the medium or high participant group on average for the course) than students who did not report plans to become teachers (approximately 50% in the medium or high participant group). The *n* for the latter group was much smaller, precluding a significant difference between average participation levels for the two groups.

The top reason for participation reported by students in Howard and Henney's (1998) article was to seek clarification of information. This reason was the only one selected by more than 50% of the sample of traditional students. In contrast, non-traditional students more strongly endorsed virtually all the listed reasons for participation than did the traditional students. Fewer

students in the Howard and Henney study reported participating because they had something to contribute to the class or because they learn by participating. Even fewer students reported they enjoyed participation, participated because it was part of their grade, were trying to make class more interesting, or felt obligated to participate on account of others' reticence.

There are several points of comparison between results from the Howard and Henney (1998) study and those from the current study. A minority of students in the Howard and Henney study reported they had something to contribute to the class, and almost half of the students (45%) in the current study reported they were uncertain as to whether they had insights that would benefit their peers. Less than half of the Howard and Henney sample reported that they learned by participating. Nearly half of the students in the current study did not report a relationship between regular or heavy participation and their learning. Interestingly, the majority of students in the current study reportedly perceived a high correlation between participation and course performance. The top reason reported by students in the Howard and Henney study (seeking clarification) was not directly assessed in the current study. It is possible that seeking clarification carries more weight for students in lecture-based classrooms than classes characterized by open discussion; student questions were much less frequent than student responses to instructor questions in the current study. Nevertheless, the results of the current study suggest that, if left up to students, seeking clarification may be less important than history and confidence of participation.

Howard et al. (2002) found that "talkers" were more likely to perceive participation as personally beneficial and helpful to the class as a whole, have greater confidence in speaking up, and have a higher comfort level than "non-talkers." The current study also showed that low and

high participants in the classroom have significantly different attitudes toward participation. For most items on the survey, the high-participant group mean was higher than the low-participant group mean, indicating a more favorable attitude about various specific aspects of participation, such as greater expectancy of participation or higher perceived value of participation.

A survey item measuring students' perception of the quantity-quality relationship of participation produced results inconsistent with those attained in a different study (Sommer & Sommer, 2007). In the current study, students had mixed views about this relationship on the beginning-of-semester survey. On one survey item, 40% of students indicated quantity is more likely to detract from quality than contribute to it, while about 30% indicated the opposite: quantity is more likely to contribute to quality than detract from it. Eighteen percent reported the two dimensions are unrelated, and a small percent of the sample reported having either extreme view (i.e., the quantity consistently contributes to, or detracts from, quality). Sommer and Sommer found that 71% of students believed credit raised the quality of class participation, while 0% reported it lowered discussion quality.

Lastly, despite mixed results found for gender in previous studies (e.g., Howard & Henney, 1998), virtually no correlation between gender and participation was found in the current study. A small but significant correlation was found between academic classification and rates of participation. Students in the advanced college years were slightly more likely to participate than those in the early years. Howard and Henney found greater participation rates among non-traditional students, suggesting that age likely has a stronger correlation with participation rate than academic classification.

Implications for Instructors

Procedural details outlined in the current study can benefit instructors by allowing for greater efficiency and objectivity in crediting students for classroom participation. Using a record-card system that also facilitates the monitoring of other course credit-producing activities (e.g., attendance, homework) increases student submission rates and provides instructors with the opportunity to check in with students daily. Instructors could expand upon the system used in the current study to daily receive other information of interest from students (e.g., general feedback, questions from discussion and points that may need clarification or elaboration) although doing so may reduce the efficiency of the current system. Thus, the record-card system can permit the collection of frequent and objective records, preferable to general ratings and likely leading to more accurate grading than the other methods, as well as greater care in participating (i.e., better quality) due to the greater effort expended in recording. Although the latter claim was not substantiated in this study, future research that better defines and assesses quality of participation may find it to be valid. Furthermore, regular self-recording and assignment of points provides frequent opportunities for students to evaluate and improve their performance.

When deciding to use the crediting procedures and/or record-card system, instructors first need to determine if the goal is to increase individual student participation or balance participation across students. Although the former was not a focus in the current study due to the cap on participation, crediting procedures have proven effective in doing so (Boniecki & Moore, 2003). Increasing student participation may be the aim in classrooms with fewer students as well as in very large lecture-based courses. Balancing discussion requires a different set of considerations, which begins with selecting an appropriate contingency level. A criterion on the

number of comments desired per student for each class session should be set to discourage any students from dominating the discussion and to motivate reticent students to become involved in discussion; the criterion can be considered as a cap for talkative students and a goal for reticent students. Additionally, weighting the first comment more heavily than subsequent comments may be a significant factor in motivating reticent students to participate.

Reducing the number of dominating participants in a class may be as challenging as reducing the number of non-participants. Hodge and Nelson (1991) reflected on the fact that “overparticipants” seemed frustrated whenever the instructor did not call on them or limited their comments. These students may have initially misperceived the goal of the intervention in their study as to increase every student’s participation. Although this misperception and resulting frustration were likely prompted by the nature of their study (a demonstration of differential reinforcement in which the dominating students were not rewarded), the importance of instructors explicitly addressing the goal of the intervention is highlighted. Specifically, students would likely benefit from instructors firmly establishing the expectation of participation on the first day of class and noting that the instructor would first recognize voluntary comments from students who had not yet contributed.

Instructional and course arrangements in the current study circumvented many problems that may occur for different class setups. Such contextual and classroom managerial issues include (1) preparing students for each discussion, (2) arranging the classroom environment for large class discussions, and (3) building adequate relationships with students. Taking some measure to prepare students for discussion—including but not limited to assigned readings or homework, study questions, small group discussions—appears to be a prerequisite for

encouraging balanced and quality participation. One idea would be to formulate a set of higher-order questions (referred to as comprehension questions in this study) per class period and make the questions available to students through email or a course web site ahead of time. Such questions could be used as the framework for discussion. Allowing access to the questions could lead to greater student preparation, more balanced participation, and higher quality participation.

Although an ideal classroom arrangement would likely involve positioning chairs in a circle, a larger class size makes this practice prohibitive. In the current study, although chairs were in rows, students were generally able to see each other because rows were elevated and arranged in a semi-circular fashion. Students were able to identify each other by name because they were written on both sides of a large namecard, which were placed on their desk. Also seeking to create a supportive classroom culture, building adequate relationships with students may be accomplished by the following strategies: (1) learning the names of students and showing personal interest, (2) listening actively and showing respect for and value of students' comments, and (3) encouraging students to visit instructors during predetermined office hours.

Limitations

Several aspects of the current study place limitations on the internal validity and generalization of the results. One threat to the validity of this study relates to the research design, which restricted the type of analyses that could be performed. Use of percent of participants as the dependent variable rather than frequency of participation limited analyses to visual inspection and proportions testing to show definitive differences between treatment and non-treatment phases.

Withdrawal of certain treatment conditions in each section of the course poses slight uncertainty in the interpretation of results. Namely, did the intervention of credit or self-recording unintentionally carry over to the subsequent phase? Removal of credit in certain phases of the study does not appear problematic, as students were clearly informed when credit for participation was unavailable; however, collecting record cards before class and not requiring self-recording does not guarantee students were not monitoring their participation. Students may have continued to tally their comments, mentally or on paper for their own records. Nevertheless, only Section A experienced removal of self-recording.

Although ANOVA results point to an uncontaminated treatment effect of contingent credit for balancing student participation in the current study, I was unable to precisely control for instructor behavior. Instructors may have differed in other important ways, beyond type and frequency of question asked and feedback given. For example, Auster and MacRone (1994) found that students participated more in courses with female instructors; however, students' increased participation related more to how often instructors engaged in certain interactional behaviors than to gender itself. Nevertheless, it is important to note that the target interactional behaviors described by the researchers (e.g., calling on students by name, providing positive feedback, asking more analytical than factual questions, and asking for students' opinions) were practices used frequently by all three instructors in the current study.

Other methodological issues, such as the presence of observers and the time limit of the class session, are possible confounding variables. Students were not specifically informed of the reason for the observers' presence; however, they likely understood the discussion was being objectively tracked, which may have contributed to the high accuracy of student recording.

Circumventing reactivity would require a covert arrangement, which presents a problem for human subjects review. Additionally, the mere fact that each class session consists of a finite amount of time limits the number of comments students can make. Setting a two-comment criterion for the 50-min period was deemed appropriate and effective in the current study. However, students desiring to comment further were not given the opportunity. In this instance, measuring the number of hands raised would be preferable to the number of comments made.

Student evaluation of the credit and self-recording arrangements was not measured at the conclusion of the semester course. A follow-up interview with randomly selected students or non-participants to assess treatment acceptability would likely provide insight that would enhance efforts to balance participation. Some students indicated they did not like recording their comments and/or receiving credit for participation. I am inclined to agree with a point made by Henning (2005): “Whether secondary school, college, or graduate level, students are students, and they tend to value what is graded” (p. 93). Yet, still debatable is the philosophical question of whether incentives should be given for an inherently valuable activity. It may be inferred from studies documenting the effect of extrinsic rewards on intrinsic motivation (Cameron, 2001; Deci, Koestner, & Ryan, 1999; Kohn, 1995; Lepper, Greene, & Nisbett, 1973) that the reward of credit may actually diminish high participants’ inclination to participate in the future, while not lead to deleterious effects for students with initially low intrinsic motivation to participate. Conducting a longitudinal study may actually show that offering credit alters the learning history of initially low- or non-participants. In other words, an interesting question is whether the incentive of credit for students with limited motivation to participate translates to increased participation in future courses.

Future Directions and Conclusions

Ezzedeen (2008) claimed that “discussion is probably one of the most powerful pedagogical tools available to teachers (despite being anxiety-provoking to both them and their students) because it exposes students to a variety of view-points and allows them to relate to others, fostering cognitive and affective competencies...” (p. 235). Increased participation may lead to better course grades, enhanced communication skills, higher-order and critical thinking skills, and many other positive outcomes yet to be identified and studied. The current study has provided a blue print for efficiently and reliably recording quantity of participation and increasing discussion through direct contingent credit. Nonetheless, much research remains to be conducted as to how to improve quality of participation, while balancing quantity of participation across students.

Students inclined to participate, presumably based on a longstanding history of frequent participation were further motivated by credit. Instructors may be prone to view their low participants as a homogenous group and, therefore, select a strategy expected to produce results across students. However, given the variety of explanations for low participation (e.g., a history of low participation in courses, discomfort speaking in public, lack of interest in the content, personal circumstances precluding preparation for class), more individualized approaches may be necessary. Certainly, several strategies have been effective in increasing the percentage of students who participate: arranging the class so that students are well prepared for participation, giving many opportunities for students to respond to comprehension questions, inviting student questions and comments about issues under discussion, providing positive feedback when students participate, and giving formal credit for participation. Unfortunately, the findings of this

study leave unresolved the issue of what teachers can do to mobilize participation from the most reticent students.

Despite using all the strategies enumerated in the previous paragraph, some students still will not participate. We are then left with a number of questions about the non-participating students: To what extent does having acquaintances or friends in the class influence participation? Would a higher level of credit be greater incentive to participate? Would it be helpful to allow students some control in selecting the weight of credit or a separate type of reward? Are these students doing themselves a disservice by not participating? If a student elects not to participate in class discussion for whatever reason, is the teacher responsible for engaging that student in discussion? Does voluntary participation make for a better learning environment than the instructor's simply calling on students to respond? Is the claim by some students that they learn better by listening than by engaging in class discussion a credible claim? Is non-participating often analogous to non-thinking about the issues under discussion? Stated differently, does participation suggest a higher level of cognitive engagement in the discussed subject matter than does non-participation? The answers to these questions will indicate whether involvement in participation is a useful goal for all students and whether researchers should continue to explore why some students elect not to engage in such discussion. In combination with cognitive performance measures, surveys such as the one used in this study may help determine the type of treatment conditions most likely to produce participation from the most reticent students.

References

- 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.
- Anecdote. (2008). In *Merriam-Webster Online Dictionary*. Retrieved August 15, 2008, from <http://www.merriam-webster.com/dictionary/anecdote>.
- 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.
- 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.
- Burchfield, C. M., & Sappington, J. (1999). Participation in classroom discussion. *Teaching of Psychology*, 26, 290-291.
- Burgess, B., & Kaya, N. (2007). Gender differences in student attitude for seating layout in college classrooms. *College Student Journal*, 41, 940-946.
- Cameron, J. (2001). Negative effects of reward on intrinsic motivation—a limited phenomenon: comment on Deci, Koestner, and Ryan (2001). *Review of Educational Research*, 71,

29-42.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum.

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.

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

Deci, E., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125, 627-668.

Ezzedeen, S. R. (2008). Facilitating class discussions around current and controversial issues. *College Teaching*, 56, 230-236.

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

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

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

- Foster, L. N., Krohn, K. R., McCleary, D. F., Miller, K. B., Nalls, M. L., Quillivan, C. C., Taylor, C. M., & Williams, R. L. (2009). Increasing low-responding students' participation in class discussion. *Journal of Behavioral Education, 18*, 173-188.
- Fritschner, L. M. (2000). Inside the undergraduate college classroom: Faculty and students differ on the meaning of student participation. *Journal of Higher Education, 71*, 342-362.
- 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.
- Garson, G. D. (2009). "Logistic Regression", from *Statnotes: Topics in Multivariate Analysis*. Retrieved 09/10/2009 from <http://faculty.chass.ncsu.edu/garson/pa765/statnote.htm>.
- Garson, G. D. (2008). "Scales and Standard Measures", from *Statnotes: Topics in Multivariate Analysis*. Retrieved 07/26/2009 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.
- Gioia, D. A. (1987). Contribution! Not participation in the OB classroom. *Organizational Behavior Teaching Review, 11*, 15-20.
- Gopinath, C. (1999). Alternatives to instructor assessment of class participation. *Journal of Education for Business, 75*, 10-14.
- Hancock, D. R. (2000). Impact of verbal praise on college students' time spent on homework. *Journal of Educational Research, 93*, 384-389.

- Hautau, B., Turner, H. C., Carroll, E., Jaspers, K., Krohn, K., Parker, M., & Williams, R. L. (2006). Differential daily writing conditions and performance on major multiple-choice exams. *Journal of Behavioral Education, 15*, 171-181.
- Henning, J. E. (2005). Leading discussions: Opening up the conversation. *College Teaching, 53*, 90-94.
- Hodge, G. K., & Nelson, N. H. (1991). Demonstrating differential reinforcement by shaping classroom participation. *Teaching of Psychology, 18*, 239-241.
- 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.
- Howard, J. R., Short, L. B., & Clark, S. M. (1996). Students' participation in the mixed-age college classroom. *Teaching Sociology, 24*, 8-24.
- 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.
- Karp, D. A., & Yoels, W. C. (1976). The college classroom: Some observations on the meaning of student participation. *Sociology and Social Research, 60*, 421-439.

- Kember, D., & Gow, L. (1994). Orientations to teaching and their effect on the quality of student learning. *Journal of Higher Education*, 65, 58-74.
- Kohn, A. (1995). Punished by rewards? A conversation with Alfie Kohn. *Educational Leadership*, 53(1), 13-16.
- Krohn, K. R., Foster, L. N., McCleary, D. F., Miller, K. B., Nalls, M. L., Quillivan, C. C., Taylor, C. M., & Williams, R. L. (2010). Reliability of students' self-recorded participation in class discussion. Manuscript submitted for publication.
- Lepper, M. R., Greene, D., & Nisbett, R. F. (1973). Undermining children's intrinsic interest with extrinsic rewards: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, 28, 129-137.
- Love, K. B. (1981). Comparison of peer assessment methods: Reliability, validity, friendship bias and user reaction. *Journal of Applied Psychology*, 66, 451-457.
- Mainkar, A. V. (2008). A student-empowered system for measuring and weighing participation in class discussion. *Journal of Management Education*, 32, 23-37.
- Marx, A., Fuhrer, U., & Hartig, T. (1999). Effects of classroom seating arrangements on children's question-asking. *Learning Environments Research*, 2, 249-263.
- Messineo, M., Gaither, G., Bott, J., & Ritchey, K. (2007). Inexperienced versus experienced students' expectations for active learning in large classes. *College Teaching*, 55, 125-133.
- McCleary, D. F., Foster, L. N., & Williams, R. L. (2010). Relationship 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.

- Montello, D. R. (1988). Classroom seating location and its effect on course achievement, participation, and attitudes. *Journal of Environmental Psychology*, 8(2), 149-157.
- Nunn, C. E. (1996). Discussion in the college classroom: Triangulating observational and survey results. *Journal of Higher Education*, 67, 243-266.
- Rosenfeld, P., Lambert, N., & Black, A. (1985). Desk arrangement effects on pupil classroom behavior. *Journal of Educational Psychology*, 77, 101-108.
- Simpson, E., & 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.
- Smith, D. H. (1992). Encouraging students' participation in large classes: A modest proposal. *Teaching Sociology*, 20, 337-339.
- Sommer, R., & Sommer, B. A. (2007). Credit for comments, comments for credit. *Teaching of Psychology*, 34, 104-106.
- Voelkl, K. E. (1995). School warmth, student participation, and achievement. *Journal of Experimental Education*, 63, 127-138.
- Weaver, R. R., & Qi, J. (2005). Classroom organization and participation: College students' perceptions. *The Journal of Higher Education*, 76, 570-601.

Appendixes

Appendix A: Tables

Table 1

Flowchart of Treatment Combinations and Sequence for Each Section

	Unit				
	1	2	3	4	5
Section A ^a	NSR-NC	NSR-NC	SR-C	NSR-NC	SR-C
Section B ^b	NSR-NC	SR-C	SR-NC	SR-C	SR-NC
Section C ^c	NSR-NC	NSR-C	NSR-NC	NSR-C	NSR-NC

Note. NSR = No self-recording; SR = Self-recording; C = Credit for participation; NC = No credit for participation.

^a the third class session meeting on Mondays, Wednesdays, and Fridays (MWF) was randomly assigned to this contingency arrangement; ^b first class session meeting on MWF; ^c second class session meeting on MWF.

Table 2

Mean Daily Participation Level and Standard Deviation Based on Student Gender and Academic Classification

	Daily Participation Mean				
	Overall	NSR-NC	NSR-C	SR-C	SR-NC
Gender					
Male	1.78 (1.44)	1.91 (1.79)	1.85 (1.36)	1.58 (1.09)	1.55 (1.20)
Female	1.46 (1.30)	1.54 (2.01)	1.68 (1.22)	1.51 (.92)	1.56 (.92)
Academic Classification					
Freshman	1.69 (1.63)	2.50 (4.35)	0.50 (-)	1.88 (.82)	1.88 (.82)
Sophomore	1.35 (1.22)	1.42 (1.73)	1.75 (1.19)	1.36 (.81)	1.38 (.88)
Junior	1.58 (1.39)	1.59 (1.82)	1.75 (1.37)	1.56 (1.01)	1.62 (.88)
Senior	1.58 (1.07)	1.38 (1.16)	1.75 (1.43)	2.05 (1.37)	1.13 (1.24)
Graduate	2.81 (1.87)	3.48 (2.87)	2.53 (1.84)	2.16 (1.61)	2.16 (1.61)

Note. Overall = the daily participation mean computed with all 5 units; NSR-NC = the daily participation mean in units with no self-recording or credit; NSR-C = the daily participation mean in units with credit but no self-recording; SR-C = the daily participation mean in units with self-recording and credit; SR-NC = the daily participation mean in units with self-recording but no credit.

Dash indicates no standard deviation is available because $n = 1$ for that condition.

Table 3

Correlations between Student and Primary Observer Records of Class Participation on Baseline (b), Credit (c), and Non-Credit (nc) Units

	Units				
	1	2	3	4	5
Section A	b	b	c	nc	c
Day 1	NSR	NSR	.86	NSR	.93
Day 2	NSR	NSR	.96	NSR	.99
Day 3	NSR	NSR	.98	NSR	.97
Day 4	NSR	NSR	.97	NSR	.96
Section B	b	c	nc	c	nc
Day 1	NSR	.96	.94	.91	.97
Day 2	NSR	.92	.93	.91	.94
Day 3	NSR	.91	.91	.98	.94
Day 4	NSR	.99	.93	.93	.96
Section C	b	c	nc	c	nc
Day 1	NSR	NSR	NSR	NSR	NSR
Day 2	NSR	NSR	NSR	NSR	NSR
Day 3	NSR	NSR	NSR	NSR	NSR
Day 4	NSR	NSR	NSR	NSR	NSR
All Sections					
Day 1	NSR	.96	.91	.91	.96
Day 2	NSR	.92	.92	.91	.95
Day 3	NSR	.91	.93	.98	.94
Day 4	NSR	.99	.94	.93	.96

Note. ^b = baseline days (no credit or self-recording), ^c = credit days, ^{nc} = non-credit days; NSR = no self-recording.

Table 4

Correlations between Student and Observer Records of Class Participation on Inter-rater Check Days

Pairs within sections	Units				
	1	2	3	4	5
Section A	(<i>n</i> = 48) ^b	(<i>n</i> = 46) ^b	(<i>n</i> = 50) ^c	(<i>n</i> = 48) ^{nc}	(<i>n</i> = 41) ^c
Students and observer 1	NSR	NSR	.98	NSR	.97
Students and observer 2	NSR	NSR	.98	NSR	.97
Observers 1 and 2	1.0	.99	1.0	.99	1.0
Section B	(<i>n</i> = 54) ^b	(<i>n</i> = 52) ^c	(<i>n</i> = 48) ^{nc}	(<i>n</i> = 49) ^c	(<i>n</i> = 45) ^{nc}
Students and observer 1	NSR	.91	.91	.98	.94
Students and observer 2	NSR	.93	.91	.98	.94
Observers 1 and 2	1.0	.99	1.0	1.0	1.0
Section C	(<i>n</i> = 52) ^b	(<i>n</i> = 51) ^c	(<i>n</i> = 50) ^{nc}	(<i>n</i> = 50) ^c	(<i>n</i> = 46) ^{nc}
Students and observer 1	NSR	NSR	NSR	NSR	NSR
Students and observer 2	NSR	NSR	NSR	NSR	NSR
Observers 1 and 2	.94	.99	.99	1.0	.99
All sections	(<i>n</i> = 154)	(<i>n</i> = 149)	(<i>n</i> = 148)	(<i>n</i> = 147)	(<i>n</i> = 132)
Student and observer 1	NSR	.91	.93	.98	.94
Student and observer 2	NSR	.93	.93	.98	.94
Observers 1 and 2	.96	.99	.99	.99	.99

Note. ^b = baseline days (no credit or self-recording), ^c = credit days, ^{nc} = non-credit days; NSR = no self-recording.

Table 5

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

Section	Units									
	Unit 1		Unit 2		Unit 3		Unit 4		Unit 5	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Section A	b		b		c		nc		c	
Student	NSR		NSR		1.36 (1.06)		NSR		1.56 (1.18)	
Observer 1	1.60 (2.20)		1.24 (2.16)		1.36 (1.10)		1.52 (2.75)		1.59 (1.34)	
Observer 2	1.60 (2.20)		1.20 (2.08)		1.36 (1.10)		1.42 (2.66)		1.59 (1.34)	
Section B	b		c		nc		c		nc	
Student	NSR		1.60 (1.01)		1.15 (1.29)		1.53 (1.02)		0.98 (1.64)	
Observer 1	2.17 (3.15)		1.69 (1.11)		1.35 (1.68)		1.59 (1.10)		1.33 (2.35)	
Observer 2	2.15 (3.13)		1.71 (1.13)		1.35 (1.68)		1.59 (1.10)		1.33 (2.35)	
Section C	b		c		nc		c		nc	
Student	NSR		NSR		NSR		NSR		NSR	
Observer 1	1.92 (2.08)		1.45 (1.43)		1.40 (2.04)		1.94 (1.48)		1.87 (2.53)	
Observer 2	1.90 (2.03)		1.43 (1.45)		1.38 (2.02)		1.94 (1.48)		1.80 (2.45)	
Combined										
Student	NSR		1.60 (1.01)		1.26 (1.18)		1.53 (1.02)		1.26 (1.31)	
Observer 1	1.91 (2.54)		1.47 (1.60)		1.37 (1.64)		1.69 (1.90)		1.60 (2.16)	
Observer 2	1.90 (2.51)		1.46 (1.58)		1.36 (1.63)		1.65 (1.86)		1.58 (2.12)	

Note. ^b = baseline days (no credit or self-recording), ^c = credit days, ^{nc} = non-credit days; NSR = no self-recording.

Table 6

Correlations between Student and Primary Observer Records of Relevant Student Participation on Baseline (b), Credit (c), and Non-Credit (nc) Units

	Units				
	1	2	3	4	5
Section A	b	b	c	nc	c
Day 1	NSR	NSR	.77	NSR	.86
Day 2	NSR	NSR	.91	NSR	.92
Day 3	NSR	NSR	.98	NSR	.90
Day 4	NSR	NSR	.98	NSR	.90
Section B	b	c	nc	c	nc
Day 1	NSR	.92	.93	.81	.99
Day 2	NSR	.93	.94	.92	.94
Day 3	NSR	.84	.91	.95	.94
Day 4	NSR	.96	.91	.90	.95
Section C	b	c	nc	c	nc
Day 1	NSR	NSR	NSR	NSR	NSR
Day 2	NSR	NSR	NSR	NSR	NSR
Day 3	NSR	NSR	NSR	NSR	NSR
Day 4	NSR	NSR	NSR	NSR	NSR
All Sections					
Day 1	NSR	.92	.88	.81	.94
Day 2	NSR	.93	.90	.92	.93
Day 3	NSR	.84	.92	.95	.92
Day 4	NSR	.96	.93	.90	.93

Note. ^b = baseline days (no credit or self-recording), ^c = credit days, ^{nc} = non-credit days; NSR = no self-recording.

Table 7

Correlations between Student and Observer Records of Relevant Student Participation on Inter-rater Check Days

Pairs within sections	Units				
	1	2	3	4	5
Section A	(<i>n</i> = 48) ^b	(<i>n</i> = 46) ^b	(<i>n</i> = 50) ^c	(<i>n</i> = 48) ^{nc}	(<i>n</i> = 41) ^c
Students and observer 1	NSR	NSR	.98	NSR	.90
Students and observer 2	NSR	NSR	.98	NSR	.90
Observers 1 and 2	1.0	.99	1.0	.99	1.0
Section B	(<i>n</i> = 54) ^b	(<i>n</i> = 52) ^c	(<i>n</i> = 48) ^{nc}	(<i>n</i> = 49) ^c	(<i>n</i> = 45) ^{nc}
Students and observer 1	NSR	.84	.91	.95	.94
Students and observer 2	NSR	.86	.91	.95	.94
Observers 1 and 2	.99	.99	1.0	1.0	1.0
Section C	(<i>n</i> = 52) ^b	(<i>n</i> = 51) ^c	(<i>n</i> = 50) ^{nc}	(<i>n</i> = 50) ^c	(<i>n</i> = 46) ^{nc}
Students and observer 1	NSR	NSR	NSR	NSR	NSR
Students and observer 2	NSR	NSR	NSR	NSR	NSR
Observers 1 and 2	.99	.99	.99	1.0	.99
All sections	(<i>n</i> = 154)	(<i>n</i> = 149)	(<i>n</i> = 148)	(<i>n</i> = 147)	(<i>n</i> = 132)
Student and observer 1	NSR	.84	.92	.95	.92
Student and observer 2	NSR	.86	.92	.95	.92
Observers 1 and 2	.99	.99	.99	.99	.99

Note. ^b = baseline days (no credit or self-recording), ^c = credit days, ^{nc} = non-credit days; NSR = no self-recording.

Table 8

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

	Units				
	1	2	3	4	5
Section A	b	b	c	nc	c
Students and observer 1*	NSR	NSR	13% (2/16)	NSR	0% (0/16)
Observers 1 and 2**	100% (4/4)	100% (3/3)	100% (1/1)	100% (4/4)	100% (4/4)
Section B	b	c	nc	c	nc
Students and observer 1	NSR	14% (3/22)	11% (1/9)	30% (8/27)	12% (2/17)
Observers 1 and 2	100% (1/1)	100% (7/7)	100% (0/0)	100% (3/3)	100% (2/2)
Section C	b	c	nc	c	nc
Students and observer 1	NSR	NSR	NSR	NSR	NSR
Observers 1 and 2	86% (6/7)	100% (0/0)	100% (3/3)	100% (5/5)	100% (0/0)

Note. ^b = baseline unit (no credit or self-recording), ^c = credit unit, ^{nc} = non-credit unit; NSR = non self-recording unit.

*Based on the number of cases in the entire four-day unit

** Based on the number of cases on only the inter-rater day in each unit

Table 9

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

Pairs within sections	Units				
	1	2	3	4	5
Section A	b	b	sr-c	nsr-nc	sr-c
Factual Questions	91%	94%	100%	87%	87%
Comprehension Questions	95%	100%	92%	91%	90%
Total Questions	99%	98%	94%	100%	100%
Section B	b	sr-c	sr-nc	sr-c	sr-nc
Factual Questions	96%	100%	92%	89 %	100%
Comprehension Questions	94%	86%	94%	96%	82%
Total Questions	95%	93%	100%	98%	90%
Section C	b	nsr-c	nsr-nc	nsr-c	nsr-nc
Factual Questions	100%	100%	100%	90%	86%
Comprehension Questions	95%	100%	100%	100%	100%
Total Questions	97%	100%	100%	97%	95%

Note. ^b = baseline units, ^c = credit units, ^{nc} = non-credit units, ^{sr} = self-record units, ^{nsr} = non self-record units.

Table 10

Correlations between Primary and Secondary Observer Records of Instructor Feedback to Students on the Inter-rater Check Day in Each Unit

Pairs within sections	Units				
	1	2	3	4	5
Section A	(<i>n</i> = 48) ^b	(<i>n</i> = 46) ^b	(<i>n</i> = 50) ^{sr-c}	(<i>n</i> = 48) ^{nsr-nc}	(<i>n</i> = 41) ^{sr-c}
Positive feedback	1.0	.99	.97	.99	.98
Negative feedback	1.0	1.0	.48	.91	.81
Section B	(<i>n</i> = 54) ^b	(<i>n</i> = 52) ^{sr-c}	(<i>n</i> = 48) ^{sr-nc}	(<i>n</i> = 49) ^{sr-c}	(<i>n</i> = 45) ^{sr-nc}
Positive feedback	.99	.95	.97	.99	.99
Negative feedback	1.0	.84	.36*	.77	.89
Section C	(<i>n</i> = 52) ^b	(<i>n</i> = 51) ^{nsr-c}	(<i>n</i> = 50) ^{nsr-nc}	(<i>n</i> = 50) ^{nsr-c}	(<i>n</i> = 46) ^{nsr-nc}
Positive feedback	.98	.98	.99	.96	.99
Negative feedback	.88	.86	.88	.41	.82
All sections	(<i>n</i> = 154)	(<i>n</i> = 149)	(<i>n</i> = 148)	(<i>n</i> = 147)	(<i>n</i> = 132)
Positive feedback	.99	.98	.98	.98	.99
Negative feedback	.93	.88	.69	.72	.78

Note. ^b = baseline days, ^c = credit days, ^{nc} = non-credit days, ^{sr} = self-record days, ^{nsr} = non self-record days.

**p* < .05; all other correlations significant at the *p* < .01 level

Table 11

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	(NSR-NC vs. SR-C)		
0 level	24/46 > 12/44; $p = .007$	23/47 > 11/46; $p = .006$	23/47 > 11/46; $p = .006$
1-2 level	12/46 < 26/44; $p = .0008$	15/47 < 28/46; $p = .003$	10/45 < 23/43; $p = .001$
3-4 level	6/46 < 7/44; $p = .35$	6/47 < 7/46; $p = .37$	5/45 < 7/43; $p = .24$
5+ level	4/46 > 0/44; $p = .02$	3/47 > 0/46; $p = .04$	4/45 > 1/43; $p = .09$
Section B	(SR-C vs. SR-NC)		
0 level	12/50 < 23/48; $p = .007$	13/52 < 21/48; $p = .024$	11/49 < 26/49; $p = .0009$
1-2 level	28/50 > 15/49; $p = .005$	26/51 > 15/50; $p = .016$	30/49 > 14/48; $p = .0008$
3-4 level	10/50 > 7/49; $p = .23$	11/51 > 7/50; $p = .16$	9/49 > 7/48; $p = .31$
5+ level	0/50 < 3/49; $p = .04$	1/51 < 5/50; $p = .04$	0/49 < 2/48; $p = .07$
Section C	(NSR-C vs. NSR-NC)		
0 level	15/50 < 26/49; $p = .009$	17/51 < 27/50; $p = .018$	13/49 < 25/48; $p = .005$
1-2 level	21/50 > 11/48; $p = .022$	22/52 > 10/48; $p = .010$	21/49 > 11/49; $p = .016$
3-4 level	11/50 > 9/48; $p = .34$	9/52 > 7/48; $p = .36$	13/49 > 7/49; $p = .07$
5+ level	3/50 < 6/48; $p = .13$	4/52 < 6/48; $p = .21$	2/49 < 6/49; $p = .07$

Note. NSR-NC = No self-recording and no credit; SR-C = self-recording and credit; SR-NC = self-recording and no credit; NSR-C = no self-recording and credit

Table 12

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

		Units				
		1	2	3	4	5
Section A	n =	23 ^b	22 ^b	22 ^{sr c}	22 ^{nsr nc}	22 ^{sr c}
Low		23	20	11	20	12
Medium		0	2	9	2	9
High`		0	0	2	0	1
Section B	n =	22 ^b	22 ^{sr c}	22 ^{sr nc}	21 ^{sr c}	22 ^{sr nc}
Low		22	10	16	9	18
Medium		0	9	5	7	3
High`		0	3	1	5	1
Section C	n =	22 ^b	22 ^{nsr c}	22 ^{nsr nc}	22 ^{nsr c}	22 ^{nsr nc}
Low		22	16	20	10	19
Medium		0	6	2	10	3
High`		0	0	0	2	0

Note. ^b = baseline units, ^c = credit units, ^{nc} = non-credit units, ^{sr} = self-record units, ^{nsr} = non self-record units.

Table 13

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

		Units				
		1	2	3	4	5
Section A	n =	11 ^b	11 ^b	11 ^{sr c}	11 ^{nsr nc}	10 ^{sr c}
Low		0	5	1	5	1
Medium		11	5	6	3	6
High`		0	1	4	3	3
Section B	n =	11 ^b	11 ^{sr c}	11 ^{sr nc}	11 ^{sr c}	11 ^{sr nc}
Low		0.	1	6	0	3
Medium		11	7	3	8	6
High`		0	3	2	3	2
Section C	n =	9 ^b	9 ^{nsr c}	9 ^{nsr nc}	9 ^{nsr c}	9 ^{nsr nc}
Low		0	2	6	0	5
Medium		9	3	0	4	2
High`		0	4	3	5	2

Note. ^b = baseline units, ^c = credit units, ^{nc} = non-credit units, ^{sr} = self-record units, ^{nsr} = non self-record units.

Table 14

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

		Units				
		1	2	3	4	5
Section A	n =	18 ^b	18 ^b	18 ^{sr c}	17 ^{nsr nc}	17 ^{sr c}
Low		0	2	0	1	0
Medium		0	6	3	6	5
High`		18	10	15	10	12
Section B	n =	21 ^b	21 ^{sr c}	21 ^{sr nc}	21 ^{sr c}	21 ^{sr nc}
Low		0	0	0	1	1
Medium		0	5	6	4	9
High`		21	16	15	16	11
Section C	n =	23 ^b	23 ^{nsr c}	23 ^{nsr nc}	23 ^{nsr c}	23 ^{nsr nc}
Low		0	0	2	2	2
Medium		0	8	6	4	4
High`		23	15	15	17	17

Note. ^b = baseline units, ^c = credit units, ^{nc} = non-credit units, ^{sr} = self-record units, ^{nsr} = non self-record units.

Table 15

*Mean Number of Specific Categories of Student Participation (Comments, Questions, Anecdotes)
per Unit in Each Section*

	Units				
	1	2	3	4	5
Section A	b	b	sr c	nsr nc	sr c
Comments	5.02	4.85	5.39	5.03	4.93
Questions	.07	.29	.14	.22	.33
Anecdotes	.57	.21	.69	.38	.57
Section B	b	sr c	sr nc	sr c	sr nc
Comments	6.80	6.19	5.73	6.29	4.41
Questions	.02	.19	.00	.29	.10
Anecdotes	.76	.11	.39	.20	.26
Section C	b	nsr c	nsr nc	nsr c	nsr nc
Comments	6.04	6.36	5.67	6.87	5.62
Questions	.33	.29	.15	.51	.15
Anecdotes	1.15	.18	.33	.13	.26
Overall					
Comments	6.01	5.89	5.59	6.13	4.96
Questions	.14	.25	.10	.35	.18
Anecdotes	.83	.16	.47	.22	.35

Note. ^b = baseline days, ^c = credit days, ^{nc} = non-credit days, ^{sr} = self-record days, ^{nsr} = non self-record days.

Table 16

Number of Relevant and Non-Relevant Comments per Unit Recorded by Students and the

Observer

	Unit 1		Unit 2		Unit 3		Unit 4		Unit 5	
	R	NR	R	NR	R	NR	R	NR	R	NR
Section A	b		b		c		nc		c	
Student	NSR		NSR		179, 6		NSR		164, 6	
Observer	183, 10		166, 17		174, 13		167, 12		159, 13	
Section B	b		c		nc		c		nc	
Student	NSR		200, 8		186, 7		182, 13		188, 9	
Observer	207, 6		191, 17		186, 5		173, 23		185, 13	
Section C	b		c		nc		c		nc	
Student	NSR		NSR		NSR		NSR		NSR	
Observer	192, 15		184, 23		190, 8		179, 16		187, 5	

Note. R = relevant comments, NR = non-relevant comments; ^b = baseline units, ^c = credit units, ^{nc} = non-credit units;

NSR = non self-record units.

Table 17

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	b	sr c	nsr nc	sr c
Factual	53	47	26	31	23
Comprehension	84	68	58	93	56
Total	137	115	84	124	79
Section B	b	sr c	sr nc	sr c	sr nc
Factual	81	50	29	45	46
Comprehension	123	77	111	106	91
Total	204	127	140	151	137
Section C	b	nsr c	nsr nc	nsr c	nsr nc
Factual	70	48	44	39	41
Comprehension	84	66	91	94	75
Total	154	114	135	133	116

Note. ^b = baseline units, ^c = credit units, ^{nc} = non-credit units, ^{sr} = self-record units, ^{nsr} = non self-record units.

Table 18

Proportion of Positive Feedback to Relevant Student Comments and Proportion of Negative Feedback to Non-relevant Student Comments in Each Unit

	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5
Section A	b	b	sr-c	nsr-nc	sr-c
Positive/Relevant	.97	.92	1.01	1.01	1.01
Negative/Non-relevant	2.00	.79	.55	1.33	.44
Section B	b	sr-c	sr-nc	sr-c	sr-nc
Positive/Relevant	.91	.92	.98	.99	.91
Negative/Non-relevant	5.00	1.63	1.00	1.11	1.50
Section C	b	nsr-c	nsr-nc	nsr-c	nsr-nc
Positive/Relevant	.89	.89	.92	1.01	.98
Negative/Non-relevant	2.73	1.14	3.80	.64	2.75
Combined					
Positive/Relevant	.92	.91	.97	1.00	.97
Negative/Non-relevant	2.92	1.20	1.47	1.00	1.29

Note. ^b = baseline units, ^c = credit units, ^{nc} = non-credit units, ^{sr} = self-record units, ^{nsr} = non self-record units.

Table 19

Mean Scores and Standard Deviations on the Participation Survey for Low-, Medium-, and High-Participating Students

	Participant Groups		
	Low	Medium	High
Total Survey ^a	165.35 (16.83) ^{†††}	181.48 (19.94)	189.57 (13.79)
Combined Factors ^b	69.48 (11.76) ^{†††}	80.52 (13.16)	87.92 (9.92)
Factor 1: History/Confidence ^c	21.54 (5.03) ^{†††}	26.54 (5.02)	30.31 (4.25)
Factor 2: Expectation ^c	24.21 (3.91) ^{†††}	28.19 (4.76)	30.13 (3.93)
Factor 3: Personal Benefits ^d	23.73 (4.86) ^{††}	25.79 (5.21)	27.36 (3.75)
Discarded Survey Items ^e	95.88 (7.76) ^{††}	100.97 (8.29)	101.65 (5.72)

Note. ^aPossible score range was 50 - 250. ^bPossible score range was 23 – 115. ^cPossible score range was 8 - 40.

^dPossible score range was 7 – 35. ^ePossible score range was 27 – 135.

^{††}Low and High, and Low and Medium groups are significantly different ($p < .05$).

^{†††}Low and High, Low and Medium, Medium and High groups are significantly different ($p < .05$).

Table 20

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

Unit	<i>B</i>	SE	Wald statistic	Odds Ratio	95% CI for Odds Ratio	
					Lower	Upper
1 ^a	.09	.02	29.53***	1.10	1.06	1.14
2 ^b	.09	.02	26.16***	1.09	1.06	1.13
3 ^c	.06	.01	21.06***	1.06	1.04	1.09
4 ^d	.05	.01	15.65***	1.05	1.03	1.08
5 ^e	.06	.01	19.48***	1.06	1.03	1.09

*** $p < .001$.

^aSignificance for the full model, $X^2(1) = 51.26$, $p < .001$, $R^2 = .33$.

^bSignificance for the full model, $X^2(1) = 44.21$, $p < .001$, $R^2 = .34$.

^cSignificance for the full model, $X^2(1) = 30.07$, $p < .001$, $R^2 = .23$.

^dSignificance for the full model, $X^2(1) = 20.33$, $p < .001$, $R^2 = .18$.

^eSignificance for the full model, $X^2(1) = 27.30$, $p < .001$, $R^2 = .22$.

Table 21

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

	Unit				
	1	2	3	4	5
Section A	b	b	sr- c	nsr -nc	sr- c
Low Participants	78.3%	88.9%	75.0%	84.6%	76.9%
High Participants	72.2%	54.5%	85.7%	61.5%	81.3%
Overall	75.6%	78.9%	81.1%	76.9%	79.3%
Section B	b	sr-c	sr-nc	sr-c	sr-nc
Low Participants	81.8%	63.6%	81.8%	20.0%	72.7%
High Participants	70.0%	85.7%	58.8%	87.0%	35.7%
Overall	76.2%	78.1%	71.8%	66.7%	58.3%
Section C	b	nsr-c	nsr-nc	nsr-c	nsr-nc
Low Participants	68.2%	77.8%	85.7%	50.0%	80.8%
High Participants	77.3%	84.2%	43.8%	95.5%	66.7%
Overall	72.7%	81.1%	70.5%	79.4%	75.0%
All Sections Combined					
Low Participants	76.1%	78.6%	74.2%	66.7%	75.4%
High Participants	70.0%	76.5%	59.3%	79.3%	62.5%
Overall	73.2%	77.6%	67.2%	73.6%	69.7%

Note. ^b = baseline units, ^c = credit units, ^{nc} = non-credit units, ^{sr} = self-record units, ^{nsr} = non self-record units.

Table 22

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

Unit Factor	<i>B</i>	SE	Wald statistic	Odds Ratio	95% CI for Odds Ratio	
					Lower	Upper
Unit 1 ^a						
History and Confidence	.27	.07	15.84**	1.32	1.15	1.51
Expectation	.22	.10	4.91*	1.24	1.03	1.50
Personal Benefits	-.03	.07	0.14	0.97	1.03	1.13
Unit 2 ^b						
History and Confidence	.24	.07	10.19**	1.27	1.10	1.47
Expectation	.12	.10	1.33	1.13	0.92	1.38
Personal Benefits	.00	.08	0.00	1.0	0.85	1.17
Unit 3 ^c						
History and Confidence	.24	.07	13.51**	1.28	1.12	1.45
Expectation	.13	.09	2.23	1.14	0.96	1.35
Personal Benefits	-.08	.07	1.38	0.92	0.81	1.06
Unit 4 ^d						
History and Confidence	.12	.06	3.88*	1.13	1.00	1.27
Expectation	.04	.07	0.41	1.05	0.91	1.20
Personal Benefits	.02	.06	0.08	0.02	0.90	1.15
Unit 5 ^e						
History and Confidence	.23	.07	10.67**	1.26	1.10	1.44
Expectation	.11	.09	1.43	1.12	0.93	1.30
Personal Benefits	-.07	.08	0.91	0.93	0.80	1.08

* $p < .05$. ** $p < .01$.^aSignificance for the full model, $X^2(3) = 76.08$, $p < .001$, $R^2 = .45$.^bSignificance for the full model, $X^2(3) = 51.47$, $p < .001$, $R^2 = .38$.^cSignificance for the full model, $X^2(3) = 52.31$, $p < .001$, $R^2 = .36$.^dSignificance for the full model, $X^2(3) = 20.98$, $p < .001$, $R^2 = .18$.^eSignificance for the full model, $X^2(3) = 43.81$, $p < .001$, $R^2 = .33$.

Table 23

Percentage of Cases Correctly Classified into Low and High Groups for the Three Survey

Factors together as a Model for Predicting Participation

	Unit				
	1	2	3	4	5
Section A	b	b	sr- c	nsr -nc	sr- c
Low Participants	82.6%	85.2%	75.0%	84.6%	84.6%
High Participants	83.3%	63.6%	90.5%	61.5%	87.5%
Overall	82.9%	78.9%	84.8%	76.9%	86.2%
Section B	b	sr-c	sr-nc	sr-c	sr-nc
Low Participants	77.3%	63.6%	81.8%	40.0%	90.9%
High Participants	75.0%	90.5%	76.5%	91.3%	71.4%
Overall	76.2%	81.3%	79.5%	75.8%	83.3%
Section C	b	nsr-c	nsr-nc	nsr-c	nsr-nc
Low Participants	81.8%	88.9%	85.7%	58.3%	88.5%
High Participants	81.8%	89.5%	50.0%	90.9%	77.8%
Overall	81.8%	89.2%	72.7%	79.4%	84.1%
All Sections Combined					
Low Participants	79.1%	75.0%	77.4%	62.5%	77.0%
High Participants	80.0%	82.4%	77.8%	77.6%	77.1%
Overall	79.5%	78.5%	77.6%	70.8%	77.1%

Note. ^b = baseline units, ^c = credit units, ^{nc} = non-credit units, ^{sr} = self-record units, ^{nsr} = non self-record units.

Table 24

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

Unit Factor	<i>B</i>	SE	Wald statistic	Odds Ratio	95% CI for Odds Ratio	
					Lower	Upper
Unit 1 ^a						
History and Confidence	.36	.06	35.76**	1.43	1.27	1.61
Unit 2 ^b						
History and Confidence	.30	.06	28.44**	1.36	1.21	1.52
Unit 3 ^c						
History and Confidence	.28	.05	30.52**	1.33	1.20	1.47
Unit 4 ^d						
History and Confidence	.16	.04	16.53**	1.17	1.09	1.27
Unit 5 ^e						
History and Confidence	.26	.05	26.02**	1.30	1.17	1.43

* $p < .05$. ** $p < .01$.

^aSignificance for the full model, $X^2(2) = 70.79$, $p < .001$, $R^2 = .43$.

^bSignificance for the full model, $X^2(1) = 50.31$, $p < .001$, $R^2 = .37$.

^cSignificance for the full model, $X^2(1) = 50.80$, $p < .001$, $R^2 = .35$.

^dSignificance for the full model, $X^2(1) = 21.02$, $p < .001$, $R^2 = .18$.

^eSignificance for the full model, $X^2(1) = 42.16$, $p < .001$, $R^2 = .32$.

Table 25

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

	Unit				
	1	2	3	4	5
Section A	b	b	sr- c	nsr -nc	sr- c
Low Participants	78.3%	81.5%	75.0%	80.8%	76.9%
High Participants	88.9%	63.6%	90.5%	62.9%	87.5%
Overall	82.9%	76.3%	84.8%	76.9%	82.8%
Section B	b	sr-c	sr-nc	sr-c	sr-nc
Low Participants	72.7%	72.7%	77.3%	30.0%	86.4%
High Participants	76.2%	90.9%	77.8%	91.7%	57.1%
Overall	74.4%	84.8%	77.5%	73.5%	75.0%
Section C	b	nsr-c	nsr-nc	nsr-c	nsr-nc
Low Participants	86.4%	77.8%	82.1%	50.0%	76.9%
High Participants	77.3%	84.2%	52.9%	87.0%	66.7%
Overall	81.8%	81.1%	71.1%	74.3%	72.7%
All Sections Combined					
Low Participants	77.6%	75.0%	77.4%	60.4%	75.4%
High Participants	78.7%	80.8%	76.8%	78.3%	75.0%
Overall	78.1%	77.8%	77.1%	70.4%	75.2%

Note. ^b = baseline units, ^c = credit units, ^{nc} = non-credit units, ^{sr} = self-record units, ^{nsr} = non self-record units.

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 relevant (R) or non-relevant (N).	
1. R () or N () --	
2. R () or N () --	
3. R () or N () --	
Over: (☐) Yes ☐ or No ☐	

Figure 1. Record card submitted daily by students to track participation.

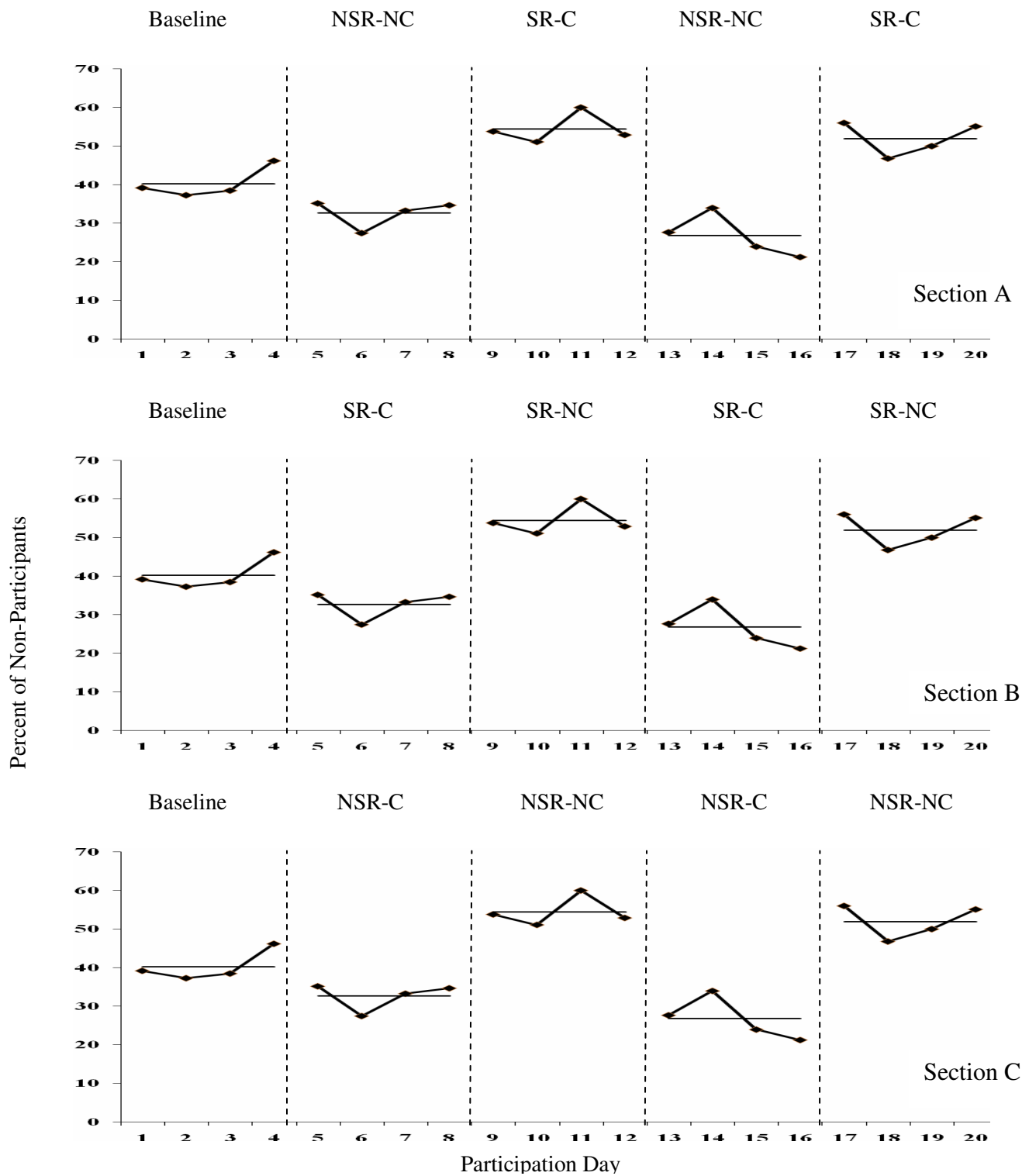


Figure 2. Percent of non-participants (0 comment level) each day.

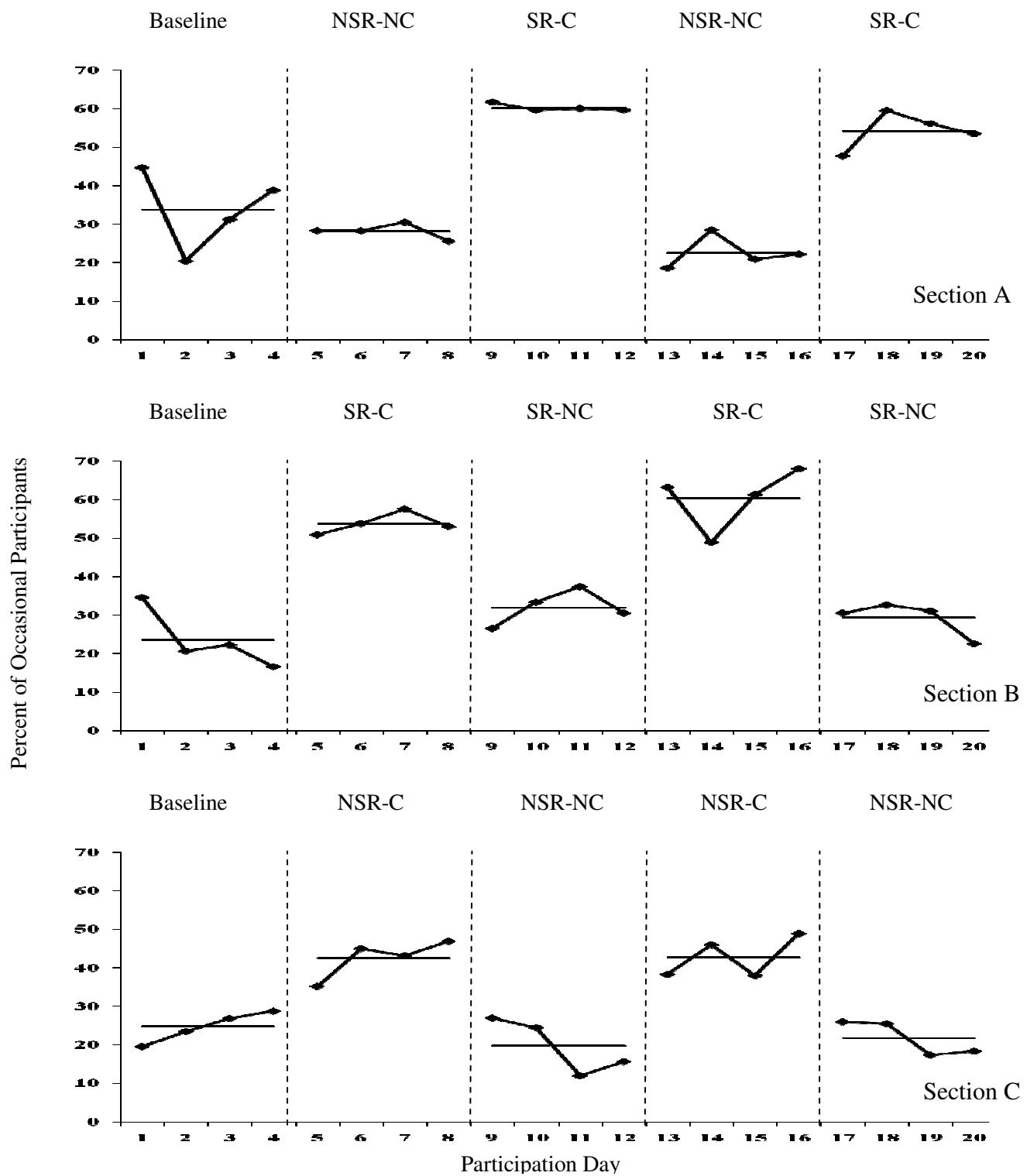


Figure 3. Percent of occasional participants (1-2 comment level) each day.

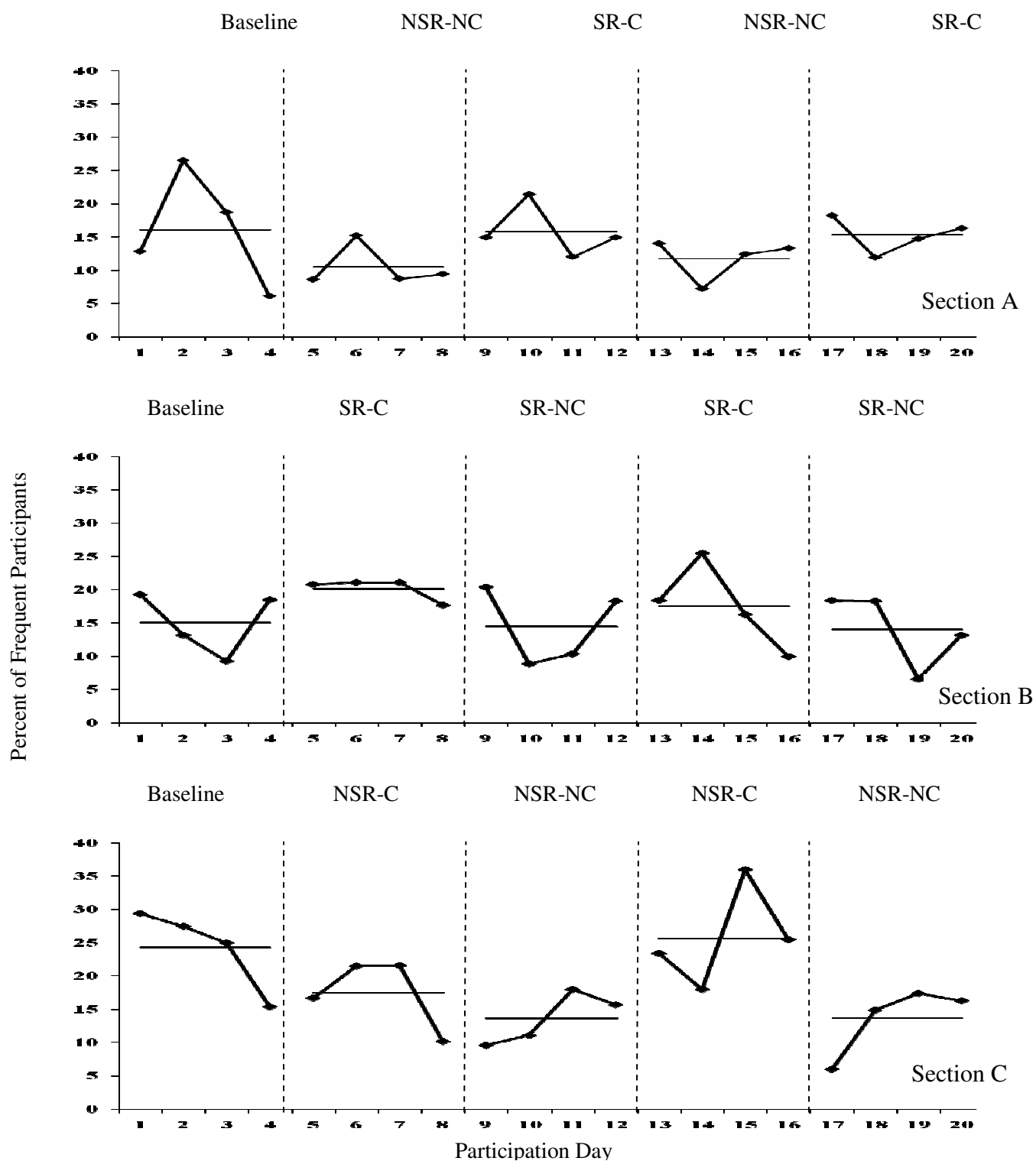


Figure 4. Percent of frequent participants (3-4 comment level) each day.

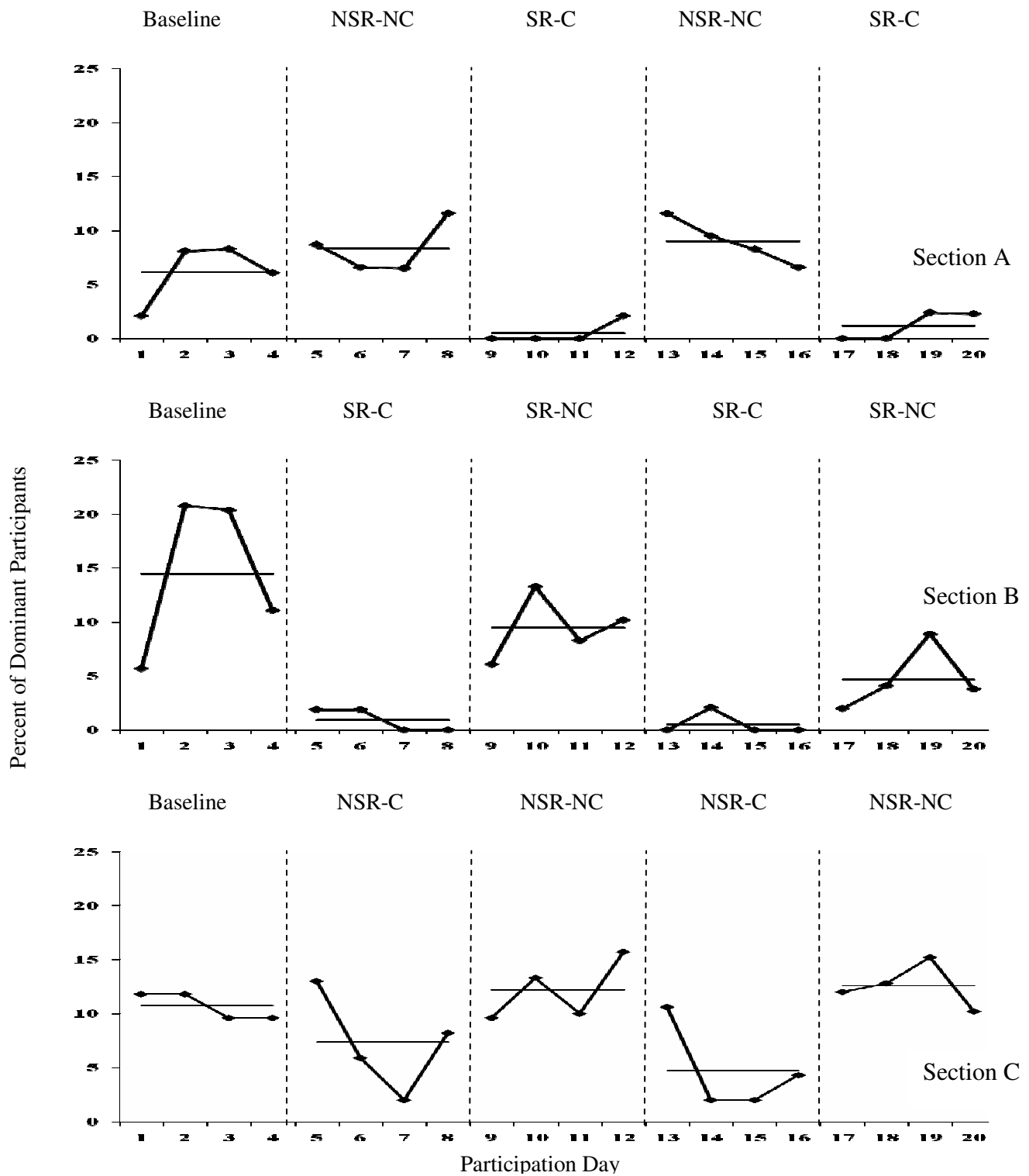


Figure 5. Percent of dominating participants (5+ comment level) each day.

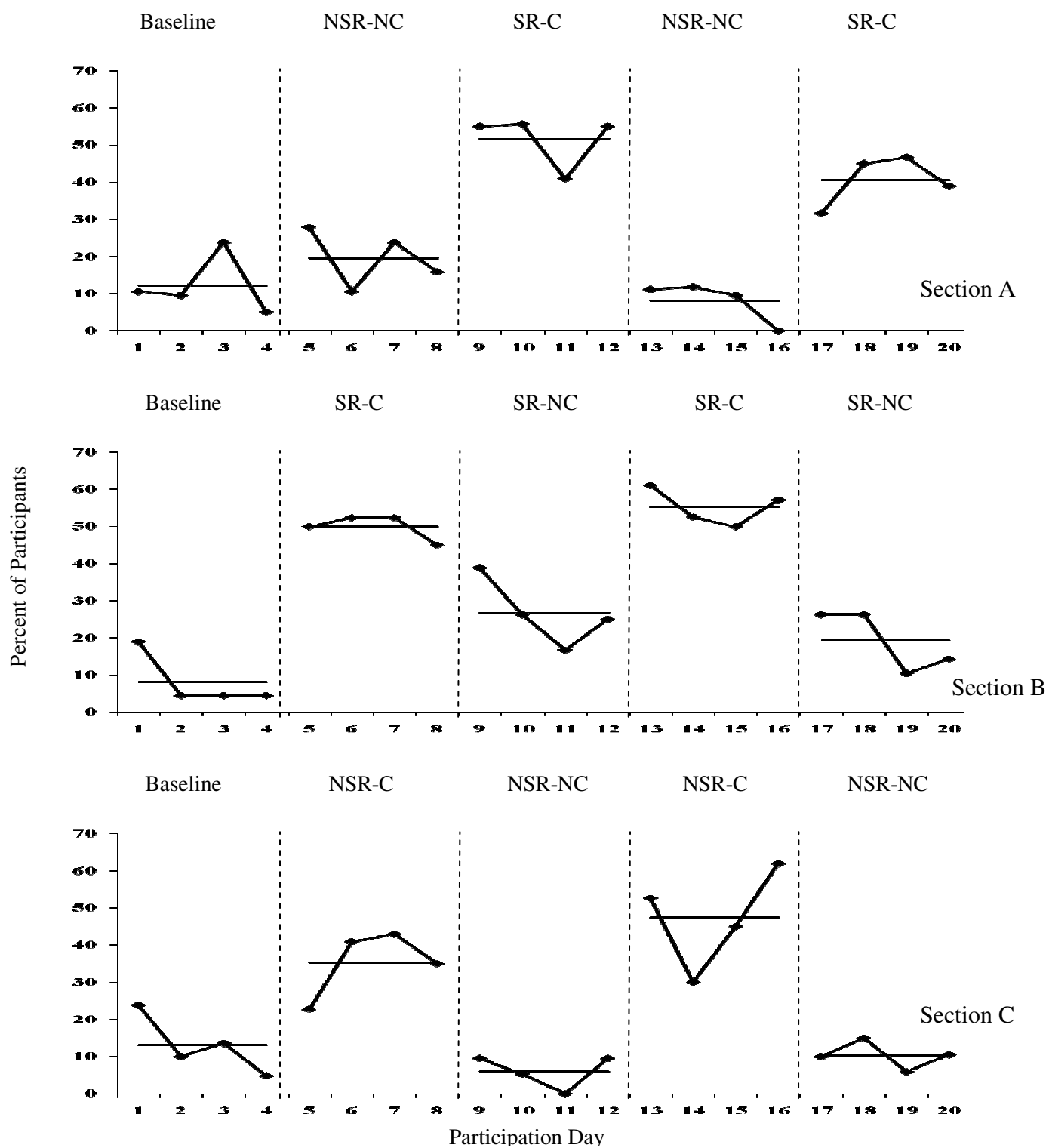


Figure 6. Percent of low participants in Unit 1 who participated in subsequent units.

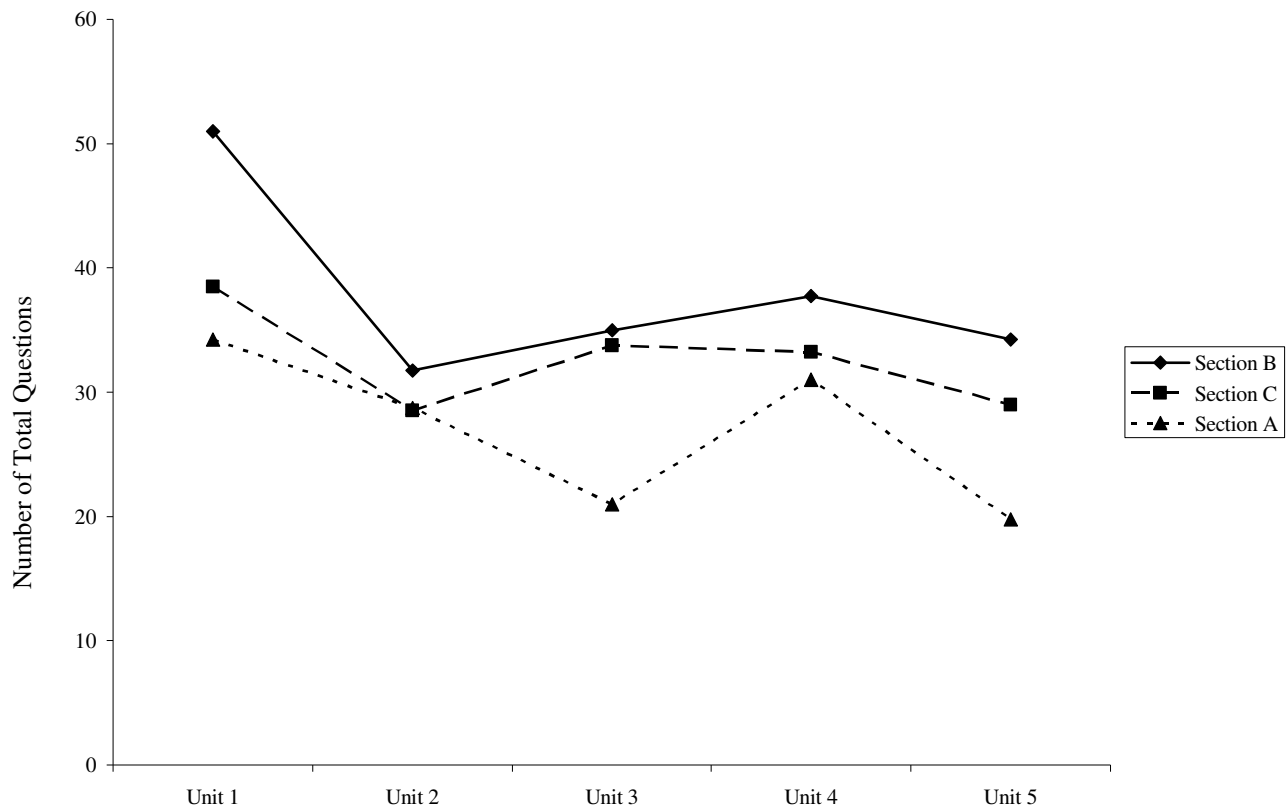


Figure 7. Mean number of total questions per unit posed by the instructor in each section.

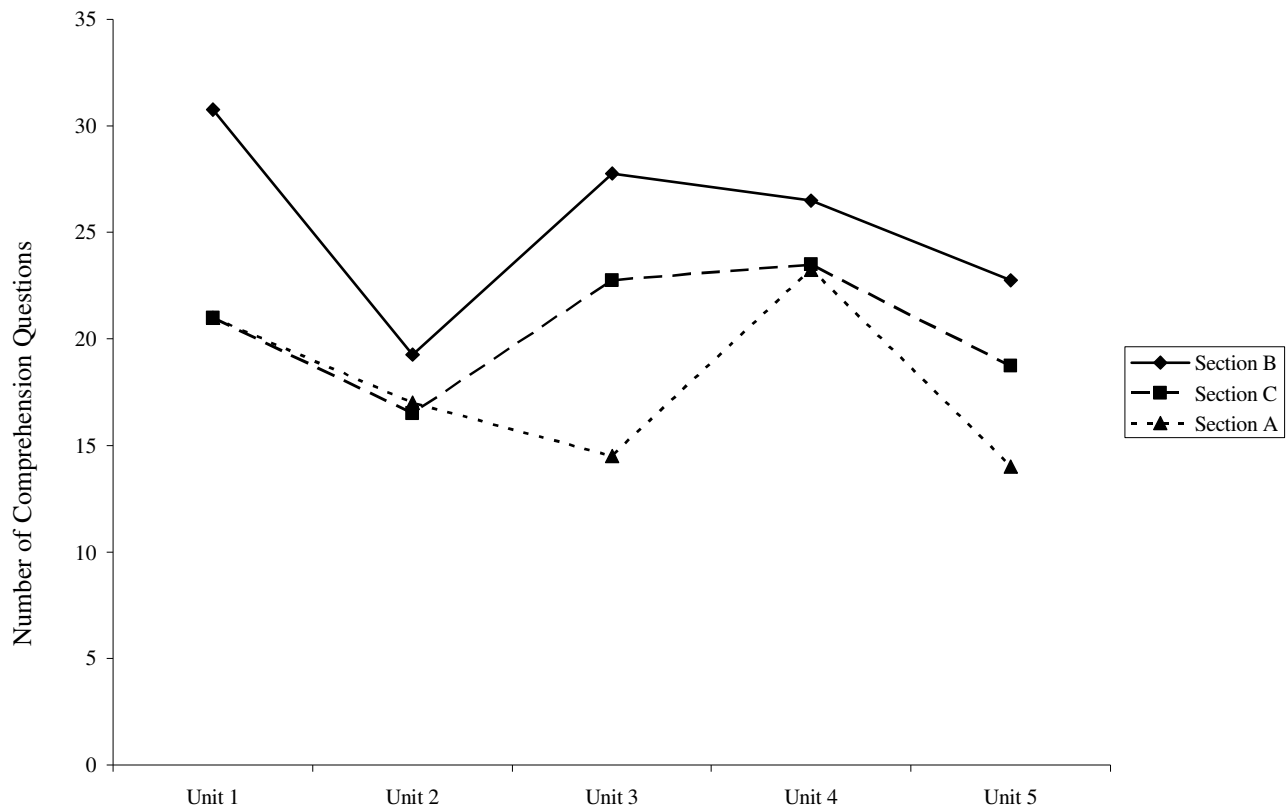


Figure 8. Mean number of comprehension questions per unit posed by the instructor in each section.

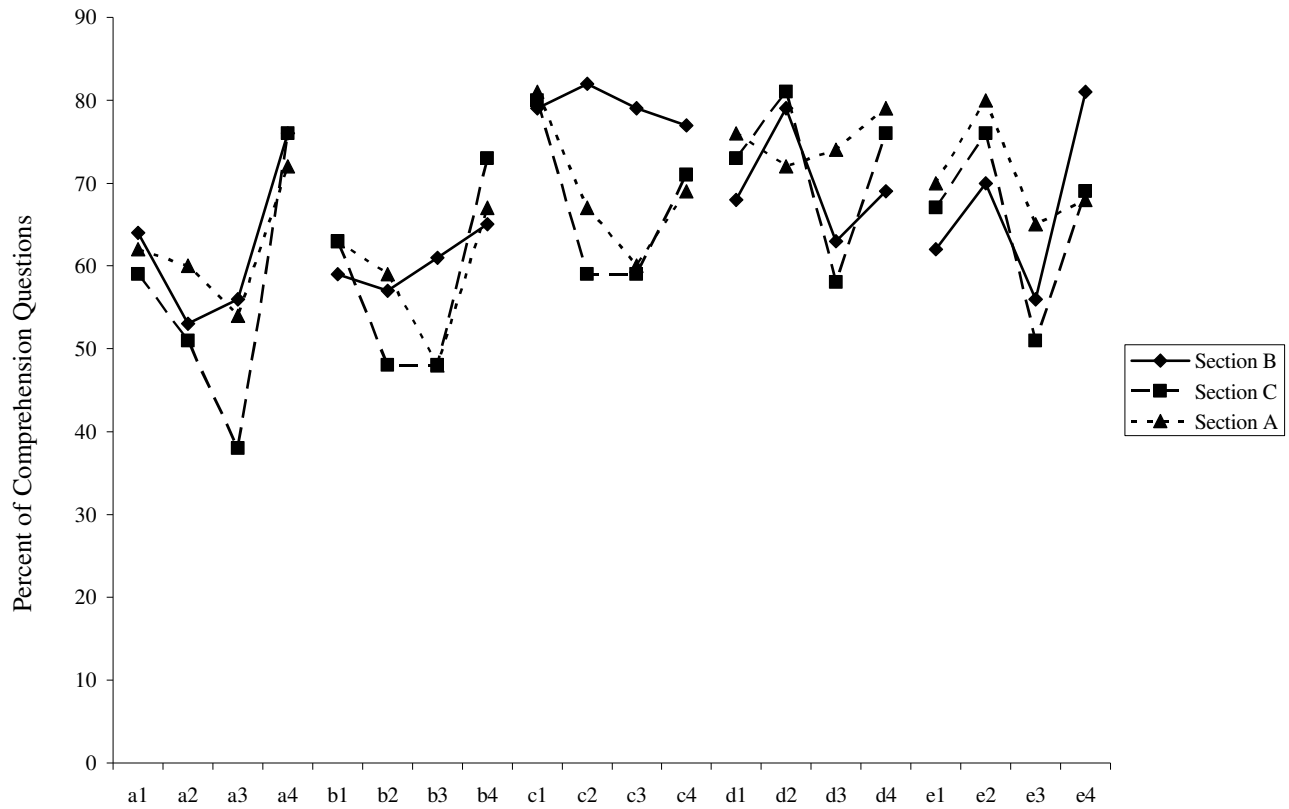


Figure 9. Percent of comprehension questions each day posed by the instructor in each section.

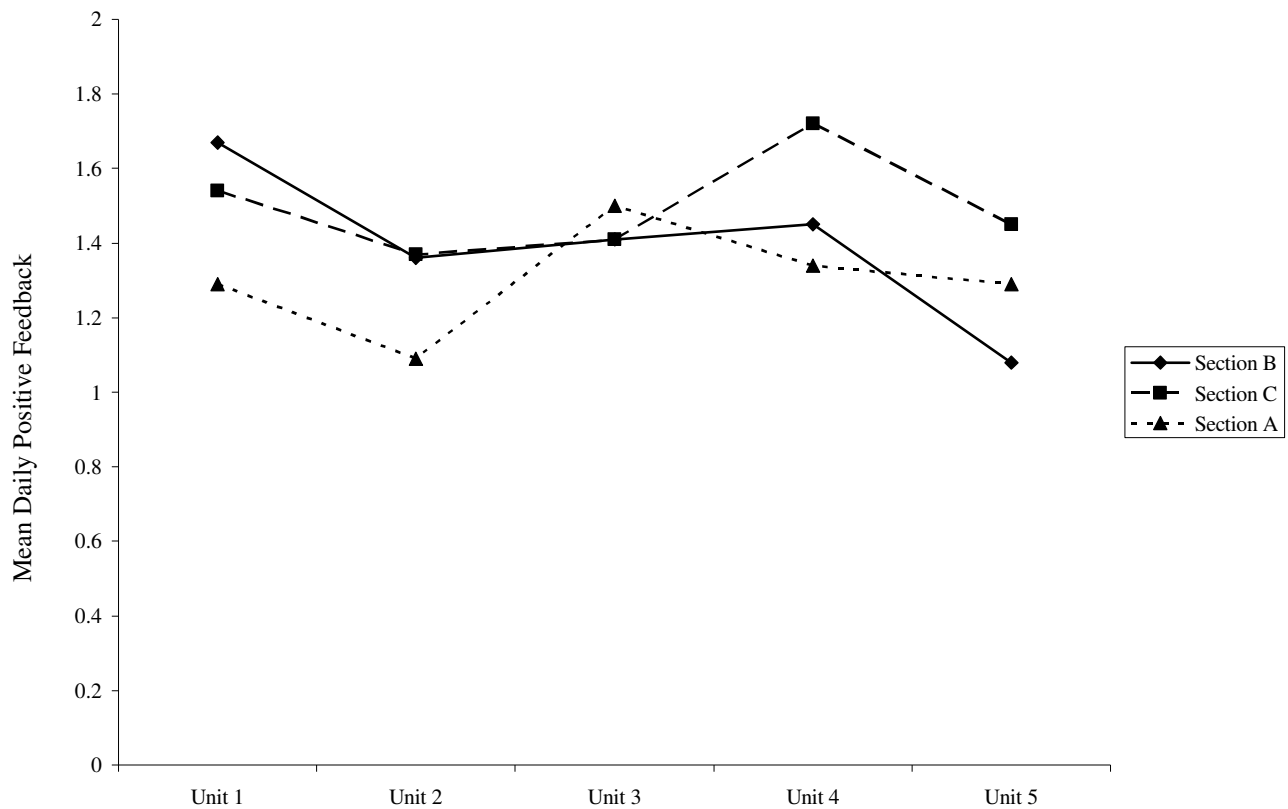


Figure 10. Mean quantity of positive feedback provided daily to individual students within each unit.

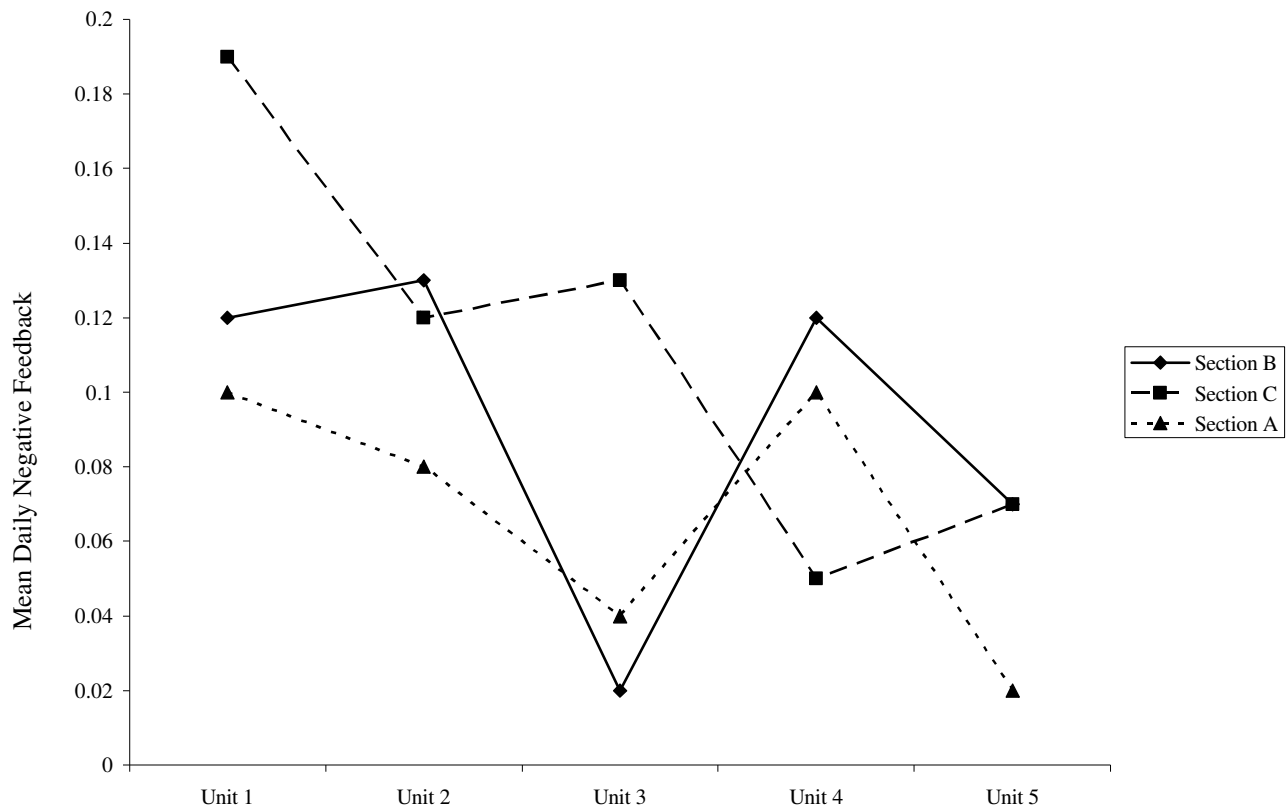


Figure 11. Mean quantity of negative feedback provided daily to individual students within each unit.

Appendix C: Instructor Discussion Form

Year_____ Semester_____ Date_____ Section_____ Instructor_____ Observer_____

[illegible]

Appendix D: Student Discussion Form

Year_____ Semester_____ Date_____ Section_____ Instructor_____ Observer_____

Relevant Criteria: Accurate, Non-redundant, On topic

[illegible]

NOTES:

Appendix E: 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? (mean = 3.40, SD = 0.84)
 - a. participating several times a day in most class discussions (16.2%)
 - b. participating once or twice a day in most class discussions (30.6%)
 - 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? (mean = 3.85, SD = 1.02)
 - a. participating several times a day in most discussions (21.2%)
 - b. participating once or twice a day in most discussions (48.8%)
 - c. participating once or twice a day in about 50% of the discussions (25.0%)
 - d. participating infrequently in discussions (3.8%)
 - e. never participating in discussions (1.2%)
3. Which of the following best describes your feelings about participating in discussions in classes as large as the 210 course? (mean = 3.40, SD = 0.99)
 - a. consistently feel comfortable when participating (12.5%)
 - b. generally feel comfortable when participating (37.5 %)
 - c. feel comfortable about half the time when participating (28.8%)
 - d. generally feel uncomfortable when participating (20.0%)
 - e. consistently feel uncomfortable when participating (1.2%)
4. Which of the following best describes how you feel when called on to participate in class discussion? (mean = 3.02, SD = 1.04)
 - a. extremely confident (3.8%)
 - b. somewhat confident (37.5 %)
 - c. neutral (21.9%)
 - d. somewhat uncomfortable (30.6%)
 - e. extremely uncomfortable (6.2%)

5. Which of the following best describes your perspective of the relevance of your comments in class discussion? (mean = 4.18, SD = 0.63)
- a. Your comments are almost always relevant. (28.8%)
 - b. Your comments are generally relevant. (62.5%)
 - c. Your comments are relevant about half the time. (6.9%)
 - d. Your comments are seldom relevant. (1.9%)
 - e. Your comments are almost never relevant. (0.0%)
6. Which of the following most accurately describes your typical level of preparation for class discussion in past courses? (mean = 3.82, SD = 0.54)
- a. completed all of the homework related to the topic to be discussed plus done some additional investigation of the topic (6.2%)
 - b. completed all of the homework related to the topic to be discussed (70.0%)
 - c. completed most of the homework related to the topic to be discussed (23.1%)
 - d. completed little of the homework related to the topic to be discussed (0.6%)
 - 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? (mean = 4.19, SD = 0.84)
- a. Students are totally responsible for volunteering comments under those circumstances. (41.9%)
 - b. Students bear most of the responsibility for volunteering comments under those circumstances. (39.4%)
 - c. Students have marginal responsibility for volunteering comments under those circumstances. (16.2%)
 - d. Students have little responsibility for volunteering comments under those circumstances. (1.2%)
 - e. Students have no responsibility for volunteering comments under those circumstances. (1.2%)
8. What is your attitude about earning course credit for participating in class discussion? (mean = 2.98, SD = 0.84)
- a. Participation should be the most heavily weighted contributor to your grade. (3.8%)
 - b. Participation should be substantially weighted in the computation of your grade. (21.2%)
 - c. Participation should be moderately weighted in the computation of your grade. (45.6%)
 - d. Participation should be minimally weighted in the computation of your grade. (28.1%)
 - e. Participation should not be considered in the computation of your grade. (1.2%)

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? (mean = 3.13, SD = 1.0)
- a. extremely positive (9.4%)
 - b. generally positive (25.6%)
 - c. neutral (36.9%)
 - d. generally negative (25.0%)
 - e. extremely negative (3.1%)
10. How do you feel when an instructor indicates at the beginning of a course that class discussion is welcomed but optional? (mean = 4.04, SD = 0.78)
- a. extremely positive (30.0 %)
 - b. generally positive (45.6%)
 - c. neutral (22.5%)
 - 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? (mean = 3.28, SD = 1.18)
- a. much more likely to ask a question (15.6%)
 - b. somewhat more likely to ask a question (29.4%)
 - c. about equally likely to ask or answer a question (32.5%)
 - d. somewhat more likely to answer a question (11.9%)
 - e. much more likely to answer a question (10.6%)
12. How would you describe the general effect of your participating in discussion on your learning in a course? (mean = 3.56, SD = 0.96)
- a. learn best when participating heavily in discussion (15.1%)
 - b. learn best when participating regularly in discussion (41.5%)
 - c. learn best when participating periodically (30.2%)
 - d. learn best when participating infrequently in discussion (10.7%)
 - e. learn best when never participating in discussion (2.5%)
13. How do you think that your keeping a record of your participation in class discussion would affect your concentration on the discussion? (mean = 3.10, SD = 1.24)
- a. would greatly increase your concentration on the discussion (16.2%)
 - b. would generally contribute to your concentration on the discussion (21.9%)
 - c. would have an uncertain effect on your concentration on the discussion (28.8%)
 - d. would generally detract from your concentration on the discussion (21.9%)
 - e. would greatly detract from your concentration on the discussion (11.2%)

14. Which of the following class formats (discussion versus lecture) do you prefer in courses you take? (mean = 2.84, SD = 0.92)
- a. all discussion (3.1%)
 - b. mainly discussion but some lecture (18.8%)
 - c. a balance between discussion and lecture (43.8%)
 - d. mainly lecture but some discussion (27.5%)
 - e. all lecture (6.9%)
15. What do you see as the relationship between the quantity and quality of class discussion? (mean = 2.92, SD = 1.09)
- a. Quantity consistently contributes to quality. (6.9%)
 - b. Quantity is more likely to contribute to quality than detract from it. (29.4%)
 - c. Quantity and quality are unrelated. (18.1%)
 - d. Quantity is more likely to detract from quality than contribute to it. (40.0%)
 - e. Quantity consistently detracts from quality. (5.6%)
16. How do you typically respond when an instructor poses a question for class discussion? (mean = 3.05, SD = 1.11)
- a. quickly speak up (3.8%)
 - b. speak up after a short delay (37.7%)
 - c. speak up but with considerable hesitancy (32.7%)
 - d. speak up only if no one else speaks up (11.3%)
 - e. not speak up even if no one else speaks up (14.5%)
17. How do you feel about a discussion format in which students volunteer comments rather than being called on by the instructor? (mean = 4.09, SD = 1.06)
- a. greatly prefer volunteering comments rather than being called on (47.5%)
 - b. somewhat prefer volunteering comments rather than being called on (25.0%)
 - c. equally comfortable with volunteering and being called on (18.1%)
 - d. somewhat prefer being called on rather than volunteering (7.5%)
 - e. greatly prefer being called on rather than volunteering (1.9%)
18. Who is responsible for a *high level of student participation* in class discussion? (mean = 3.52, SD = 0.78)
- a. exclusively the students (9.4%)
 - b. primarily the students (41.2%)
 - c. shared equally between the students and the instructor (42.5%)
 - d. primarily the instructor (6.2%)
 - e. exclusively the instructor (0.6%)

19. Who is responsible for *very limited student participation* in class discussion? (mean = 3.36, SD = 0.90)
- exclusively the students (11.9%)
 - primarily the students (28.8%)
 - shared equally between the students and the instructor (43.8%)
 - primarily the instructor (15.0%)
 - exclusively the instructor (0.6%)
20. How would a class with frequent discussion affect your evaluation of the course? (mean = 3.56, SD = 0.90)
- greatly increase your evaluation of the course (15.6%)
 - generally increase your evaluation of the course (35.6%)
 - have little effect on your evaluation of the course (39.4%)
 - generally decrease your evaluation of the course (8.1%)
 - greatly decrease your evaluation of the course (1.2%)
21. What effect does frequent discussion by other students have on your concentration in class? (mean = 3.45, SD = 1.01)
- greatly increases your concentration (13.1%)
 - moderately increases your concentration (39.4%)
 - minimally affects your concentration (31.9%)
 - moderately decreases your concentration (10.6%)
 - greatly decreases your concentration (5.0%)
22. What effect does your personal participation in class discussion have on your concentration in class? (mean = 3.93, SD = 1.05)
- greatly increases your concentration (35.6%)
 - moderately increases your concentration (34.4%)
 - minimally affects your concentration (20.6%)
 - moderately decreases your concentration (6.2%)
 - greatly decreases your concentration (3.1%)
23. What effect does the option of volunteering comments whenever you wish have on your concentration in class? (mean = 3.65, SD = 0.86)
- greatly increases your concentration (16.2%)
 - moderately increases your concentration (40.6%)
 - minimally affects your concentration (36.2%)
 - moderately decreases your concentration (5.6%)
 - greatly decreases your concentration (1.2%)

24. How does the possibility that you might be called on to respond to an instructor question affect your concentration in class? (mean = 3.71, SD = 1.14)
- a. greatly increases your concentration (30.6%)
 - b. moderately increases your concentration (30.0%)
 - c. minimally affects your concentration (23.1%)
 - d. moderately decreases your concentration (12.5%)
 - e. greatly decreases your concentration (3.8%)
25. How does frequent discussion in the class as a whole affect your enjoyment of a class? (mean = 3.76, SD = 1.10)
- a. makes the class much more enjoyable (29.4%)
 - b. makes the class somewhat more enjoyable (35.6%)
 - c. doesn't affect your enjoyment one way or the other (20.0%)
 - d. makes the class somewhat less enjoyable (11.9%)
 - e. makes the class much less enjoyable (3.1%)
26. How do you feel toward students who frequently comment in class discussion? (mean = 3.14, SD = 1.01)
- a. You greatly appreciate their frequent participation. (8.8%)
 - b. You generally appreciate their frequent participation. (26.9%)
 - c. You feel neutral toward their frequent participation. (40.0%)
 - d. You are generally annoyed by their frequent participation. (18.8%)
 - e. You are greatly annoyed by their frequent participation. (5.6%)
27. What are your academic expectations of students who frequently participate in class? (mean = 4.31, SD = 0.75)
- a. You expect them to do well in the course. (45.6%)
 - b. You expect them to do somewhat better than average in the course. (41.9%)
 - c. You expect their frequent contributions to be unrelated to their performance in the course. (11.2%)
 - 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. (0.6%)

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? (mean = 4.52, SD = 0.77)
- a. You have the strongest inclination to comment on topics about which you have the most knowledge. (63.1%)
 - b. You are moderately inclined to comment on topics about which you have the most knowledge. (29.4%)
 - c. Your knowledge about topics has little effect on your tendency to comment on those topics. (5.6%)
 - d. You feel somewhat less need to comment on topics about which you have the most knowledge. (0.0%)
 - e. You feel the least need to comment on topics about which you have the most knowledge. (1.9%)
29. To what degree does student sharing of personal experiences in class discussion contribute to the quality of the discussion? (mean = 3.74, SD = 0.86)
- a. greatly heightens the quality of class discussion (18.2%)
 - b. moderately heightens the quality of class discussion (44.7%)
 - c. has a neutral impact on the quality of class discussion (30.8%)
 - d. moderately diminishes the quality of class discussion (5.0%)
 - e. greatly diminishes the quality of class discussion (1.3%)
30. How do you typically respond when no one else is responding to a teacher question? (mean = 3.61, SD = 1.08)
- a. Attempt to answer the question when no one else is responding. (20.6%)
 - b. Wait until the silence has become somewhat uncomfortable to you before attempting to answer the question. (45.0%)
 - c. Wait until the silence has become extremely uncomfortable to you before attempting to answer the question. (8.8%)
 - d. Respond only if the instructor calls on you to answer the question. (25.6%)
 - e. Decline to respond to the question even if the instructor calls on you. (0.0%)
31. Do you believe you have insights about course concepts that would benefit your peers if you shared them in class? (mean = 3.47, SD = 0.76)
- a. definitely "yes" (7.5%)
 - b. generally "yes" (40.0%)
 - c. uncertain (45.0%)
 - d. generally "no" (6.9%)
 - e. definitely "no" (0.6%)

32. What effect do long pauses between teacher questions and student responses have on your desire to participate in class discussion? (mean = 3.04, SD = 1.12)
- a. greatly increases your desire to participate (10.6%)
 - b. moderately increases your desire to participate (25.6%)
 - c. minimally affects your desire to participate (28.9%)
 - d. moderately decreases your desire to participate (27.5%)
 - e. greatly decreases your desire to participate (7.5%)
33. How would teachers in your past college courses most likely characterize your level of participation in class discussion? (mean = 3.11, SD = 0.88)
- a. the most talkative student in class (3.1%)
 - b. among the more talkative students in class (33.8%)
 - c. talkative to an average level (36.2%)
 - d. among the less talkative students in class (25.0%)
 - e. the least talkative student in class (1.9%)
34. How would you characterize teacher views regarding the inclusion of class discussion in student grades in your past college courses? (mean = 3.01, SD = 0.77)
- a. Participation is the most important part of a student's grade. (3.1%)
 - b. Participation is among the more important contributors to a student's grade. (20.0%)
 - c. Participation is on par with several other contributors to a student's grade. (52.5%)
 - d. Participation is among the less important contributors to a student's grade. (23.8%)
 - e. Participation is not included in a student's grade. (0.6%)
35. In comparison to other classes you are taking this semester, what expectation do you have for your participation in 210 class discussion? (mean = 3.74, SD = 0.79)
- a. more participation in 210 discussion than in any other class (16.2%)
 - b. more participation in 210 discussion than in most other classes (46.9%)
 - c. about the same level of participation in 210 discussion as in other classes (31.9%)
 - d. less participation in 210 discussion than in most other courses (5.0%)
 - e. less participation in 210 discussion than in any other course (0.0%)

36. Which of the following best expresses your view of the long-term value of learning to express one's views in public? (mean = 3.89, SD = 0.85)
- a. Learning to express one's views in public is among the most important skills one can develop in school. (23.1%)
 - b. Learning to express one's views in public is among the more important skills one can develop in school. (49.4%)
 - 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. (21.9%)
 - d. Learning to express one's views in public is among the lesser skills one can develop in school. (4.4%)
 - e. Learning to express one's view in public is among the least important skills one can develop in school. (1.2%)
37. Your interpretation of how teachers feel about class discussion is best reflected in which of the following claims? (mean = 4.25, SD = 0.80)
- a. Most teachers strongly value class discussion. (43.8%)
 - b. Most teachers moderately value class discussion. (41.2%)
 - c. Most teachers are neutral toward class discussion. (11.2%)
 - d. Most teachers moderately devalue class discussion. (3.8%)
 - e. Most teachers strongly devalue class discussion. (0.0%)
38. How would most of your high school teachers likely describe your participation in class? (mean = 3.71, SD = 1.06)
- a. extremely verbal in class (23.8%)
 - b. generally verbal in class (42.5%)
 - c. occasionally verbal in class (18.1%)
 - d. generally quiet in class (12.5%)
 - e. extremely quiet in class (3.1%)
39. Many teachers try to stimulate class discussion by asking questions. Which of the following best expresses your view of most teacher questions? (mean = 3.52, SD = 0.65)
- a. Most are highly challenging. (5.0%)
 - b. Most are moderately challenging. (45.0%)
 - c. Most are routine in nature. (47.5%)
 - d. Most provide little challenge. (1.9%)
 - e. Few provide any challenge at all. (0.6%)

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? (mean = 3.41, SD = 0.64)
- a. You would feel your classmates strongly valued your comments. (3.1%)
 - b. You would feel your classmates moderately valued your comments. (40.0%)
 - c. You would feel that your classmates were neutral toward your comments. (51.9%)
 - d. You would feel that your classmates moderately devalued your comments. (5.0%)
 - 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? (mean = 4.33, SD = 0.81)
- a. You can determine whether a comment will be relevant even before you make the comment. (51.2%)
 - b. You have your first sense of whether a comment is relevant as you are making the comment. (33.8%)
 - c. You can tell whether a comment is relevant only by the instructor's reaction to the comment. (11.9%)
 - d. You can only judge the relevance of your comment when you have time to reflect on it after class. (3.1%)
 - 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? (mean = 3.95, SD = 0.78)
- a. Consistently raise your grades. (24.4%)
 - b. Generally raise your grades. (49.4%)
 - c. Have little effect on your grades. (23.1%)
 - d. Generally lower your grades. (3.1%)
 - e. Consistently lower your grades. (0.0%)
43. How do you typically feel when you have volunteered a comment in class discussion? (mean = 3.48, SD = 0.84)
- a. You feel very important in the class. (5.6%)
 - b. You feel as if you have gained some positive recognition. (52.5%)
 - c. You feel neutral about your comment. (27.5%)
 - d. You fear that you might have said the wrong thing. (13.1%)
 - e. You believe your comment has been poorly received. (1.2%)

44. How much of a personal priority is improving the amount and/or quality of your participation in class discussion? (mean = 3.45, SD = 0.83)
- a. It is your top priority. (5.6%)
 - b. It is among your highest priorities. (46.2%)
 - c. You are neutral about the prospect of improving your class participation. (39.4%)
 - d. It is among your lowest priorities. (5.0%)
 - e. It is a non-priority for you. (3.8%)
45. What is your opinion of the social status of students who participate frequently in class discussion? (mean = 3.22, SD = 0.72)
- a. They tend to be the most popular students in class. (3.8%)
 - b. They are among the more popular students in class. (26.2%)
 - c. Frequent participation has little effect on one's standing with peers. (59.4%)
 - d. They are among the less popular students in class. (9.4%)
 - e. They tend to be the least popular students in class. (1.2%)
46. What is your opinion of the social status of students who participate little, if at all, in class discussion? (mean = 2.78, SD = 0.65)
- a. They are greatly admired for their quietness. (1.9%)
 - b. They are generally admired for their quietness. (5.6%)
 - c. Their minimal participation has little effect on how peers regard them. (62.5%)
 - d. They are generally discredited for their quietness. (28.8%)
 - e. They are greatly discredited for their quietness. (1.2%)
47. What is your view of the relationship between the perceived relevance of course content and student inclination to participate in class discussion? (mean = 4.28, SD = 0.75)
- a. Students feel the greatest desire to participate in courses they view as highly relevant. (45.0%)
 - b. Students feel a moderate desire to participate in courses they view as relevant. (38.1%)
 - c. Student inclination to participate is not affected by the perceived relevance of the course content. (16.2%)
 - 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. (0.0%)

48. Which of the following best represents how students' participation in class discussion will affect their personal standing with teachers? (mean = 4.24, SD = 0.63)
- a. Students who participate frequently usually are the most liked by their teachers. (32.5%)
 - b. Students who participate frequently increase their chances of being liked by their teachers. (60.6%)
 - c. Frequent participation has little effect on how much teachers like a student. (5.0%)
 - 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? (mean = 4.12, SD = 0.69)
- a. Teacher friendliness is the number one contributor to student participation in class discussion. (28.1%)
 - b. Teacher friendliness is among the more important contributors to student participation in class discussion. (58.8%)
 - c. Teacher friendliness has little to do with student participation in class discussion. (10.6%)
 - d. Teacher friendliness is among the less important contributors to student participation in class discussion. (2.5%)
 - 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. (mean = 3.22, SD = 0.93)
- a. Teacher knowledge is a far greater contributor than teacher friendliness to class discussion. (9.4%)
 - b. Teacher knowledge is a somewhat stronger contributor than teacher friendliness to class discussion. (25.6%)
 - c. Teacher knowledge and teacher friendliness have an equal impact on class discussion. (46.2%)
 - d. Teacher friendliness is a somewhat stronger contributor than teacher knowledge to class discussion. (15.6%)
 - e. Teacher friendliness is a far greater contributor than teacher knowledge to class discussion. (3.1%)

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

	Non-Credit Days									Credit Days							
	1	2	3	4	5	6	7	8		1	2	3	4	5	6	7	8
Student #																	
(section B)																	
1	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=
2	=	=	=	=	=	=	=	=		>p	>p	=	=	=	=	=	=
3	=	<	<	=	=	<	<	=		<	=	<	=	=	<	=	<n
4	=	<	<	<	=	=	<	=		=	=	<	=	<	<	=	=
5	=	<	=	<	=	=	=	=		=	=	=	=	=	<	=	=
6	=	<	<	=	=	=	<	=		=	<	=	=	=	=	=	=
7	=	=	=	=	=	<	=	=		=	>	<	=	=	>p	<	=
8	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=
9	=	<	=	=	=	=	<	=		=	=	=	=	=	=	=	=
10	=	=	=	=	=	=	=	=		=	<	=	<	=	<	=	=
11	=	=	=	<	=	=	=	=		=	=	=	=	=	=	=	=
12	=	=	=	=	=	<	=	=		=	=	<	=	=	=	=	>
13	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=
14	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=
15	=	=	=	=	=	=	=	=		=	=	=	=	=	>p	=	=
16	=	=	=	=	=	=	=	=		=	=	=	=	=	>p	=	>p

	Non-Credit Days									Credit Days							
	1	2	3	4	5	6	7	8		1	2	3	4	5	6	7	8
17	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=
18	=	=	=	=	=	=	=	=		=	=	=	=	>p	=	=	=
19	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=
20	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=
21	=	=	=	=	=	=	=	=		>p	=	=	=	=	=	=	=
22	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	<
23	=	=	=	=	=	<	<	=		=	=	=	=	<	=	<	=
24	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=
25	=	<	=	<	=	=	=	=		=	=	=	=	=	=	=	=
26	=	=	=	=	<	=	<	=		=	=	=	=	=	<	=	=
27	<	=	=	=	=	=	=	<		=	=	=	=	=	=	=	=
28	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=
29	=	=	=	=	=	=	=	=		=	=	>p	=	=	=	=	=
30	=	=	=	=	=	=	<	=		>p	=	=	=	=	<	=	=
31	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=
32	=	=	=	=	=	=	=	>		=	=	=	=	=	=	=	=
33	=	=	=	=	=	=	=	<		=	=	=	=	=	=	=	=
34	=	=	=	=	=	=	=	=		<n	<	=	=	=	>p	=	=
35	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=
36	=	=	=	<	<	=	=	<		=	<	<	=	<	<	=	=

	Non-Credit Days									Credit Days								
	1	2	3	4	5	6	7	8		1	2	3	4	5	6	7	8	
37	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=	
38	=	=	=	=	=	=	=	=		=	=	=	=	=	>p	=	=	
39	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=	
40	=	=	=	=	=	<	<	=		=	=	=	=	=	=	=	=	
41	=	=	=	=	=	=	=	=		=	=	>p	=	=	=	=	=	
42	=	=	=	=	=	=	=	=		=	=	>	=	>p	=	=	=	
43	=	<	=	<	=	=	=	<		<	<	=	=	>p	<	<	=	
44	=	<	=	=	=	<	=	<		=	<	=	=	=	=	=	>	
45	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=	
46	=	=	=	=	=	=	=	=		=	=	<n	=	=	=	=	=	
47	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=	
48	<	<	<	=	<	=	=	=		=	=	=	=	=	=	=	=	
49	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=	
50	=	=	=	=	=	=	=	=		=	>p	=	=	=	=	=	=	
51	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=	
52	<	=	<	=	=	<	=	=		<	=	<	<	<	=	=	=	
53	=	=	=	=	<	=	=	=		=	=	<n	=	=	=	=	=	
54	=	=	=	=	=	=	=	=		=	=	=	=	=	=	=	=	
Student #																		
(section A)																		

	Non-Credit Days									Credit Days							
	1	2	3	4	5	6	7	8		1	2	3	4	5	6	7	8
55	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
56	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
57	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
58	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
59	na	na	na	na	na	na	na	na		=	=	=	=	=	=	<	<
60	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
61	na	na	na	na	na	na	na	na		<n	=	=	=	=	=	=	=
62	na	na	na	na	na	na	na	na		=	=	=	=	<	=	=	=
63	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
64	na	na	na	na	na	na	na	na		=	<	=	=	=	=	=	=
65	na	na	na	na	na	na	na	na		=	=	=	<n	=	=	=	=
66	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
67	na	na	na	na	na	na	na	na		<	=	=	=	<	=	=	=
68	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
69	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
70	na	na	na	na	na	na	na	na		=	>	=	=	<	=	=	=
71	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	>p
72	na	na	na	na	na	na	na	na		=	=	=	<	=	=	=	=
73	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
74	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=

	Non-Credit Days									Credit Days							
	1	2	3	4	5	6	7	8		1	2	3	4	5	6	7	8
75	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
76	na	na	na	na	na	na	na	na		=	=	=	=	>p	=	=	=
77	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
78	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
79	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
80	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
81	na	na	na	na	na	na	na	na		=	<	=	=	=	=	=	=
82	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
83	na	na	na	na	na	na	na	na		=	=	=	=	<	=	<	=
84	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
85	na	na	na	na	na	na	na	na		<	=	=	=	=	=	<	<
86	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
87	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
88	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
89	na	na	na	na	na	na	na	na		=	=	=	=	=	<	=	>
90	na	na	na	na	na	na	na	na		>p	>p	=	=	>p	=	=	=
91	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
92	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
93	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
94	na	na	na	na	na	na	na	na		>f	=	>p	=	>p	=	>p	=

	Non-Credit Days									Credit Days							
	1	2	3	4	5	6	7	8		1	2	3	4	5	6	7	8
95	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
96	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
97	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
98	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
99	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
100	na	na	na	na	na	na	na	na		=	=	=	=	=	=	>p	=
101	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
102	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
103	na	na	na	na	na	na	na	na		=	=	<	=	=	=	=	=
104	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	=
105	na	na	na	na	na	na	na	na		>p	<	=	=	>p	=	=	=
106	na	na	na	na	na	na	na	na		=	=	=	=	=	=	=	>p
Total* =	51	45	49	48	50	47	46	49		93	92	93	102	91	92	98	97
Total >	0	0	0	0	0	0	0	1		0	2	1	0	0	0	0	3
Total >p	na	na	na	na	na	na	na	na		5	3	3	0	7	5	2	3
Total >f	na	na	na	na	na	na	na	na		1	0	0	0	0	0	0	0
Total <	3	9	5	6	4	7	8	5		5	9	7	3	8	9	6	3
Total <n	na	na	na	na	na	na	na	na		2	0	2	1	0	0	0	1

Note: > indicates student over-reporting; < indicates student under-reporting; = 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.

*The n for each non-credit day is 54, and the n for each credit day is 106. Self-recording in non-credit units for Section A is not applicable because students only self-recorded their participation during the two credit units (4 days each). Section C is excluded from this Appendix altogether because students did not self-record their participation.

Appendix G: Survey Items Comprising Factors, Factor Loadings, and

Means for Low, Medium, and High Participants

Items on Factor 1-

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? (.530)	2.73	3.37	4.02 ^{†††}
3. Which of the following best describes your feelings about participating in discussion in classes as large as the 210 course? (.787)	2.71	3.53	3.88 ^{††}
4. Which of the following best describes how you feel when called on to participate in class discussion? (.790)	2.40	3.10	3.48 ^{††}
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? (.563)	2.60	3.14	3.60 ^{†††}
16. How do you typically respond when an instructor poses a question for class discussion? (.555)	2.33	3.02	3.75 ^{†††}
31. Do you believe you have insights about course concepts that would benefit your peers if you shared them in class? (.645)	3.17	3.51	3.71 [†]
33. How would teachers in your past college courses most likely characterize your level of participation in class discussion? (.717)	2.50	3.03	3.75 ^{†††}

38. How would most of your high school teachers likely describe your participation in class? (.498)	3.10	3.85	4.12 ^{††}
Factor 1 Total	21.54	26.54	30.31
Items on Factor 2-	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? (.483)	3.29	3.88	4.33 ^{†††}
8. What is your attitude about earning course credit for participating in class discussion? (.783)	2.48	3.07	3.35 ^{††}
12. How would you describe the general effect of your participating in discussion on your learning in a course? (.596)	2.90	3.69	4.02 ^{††}
14. Which of the following class formats (discussion versus lecture) do you prefer in courses you take? (.582)	2.33	2.88	3.23 ^{††}
20. How would a class with frequent discussion affect your evaluation of the course? (.515)	3.23	3.64	3.79 ^{††}
34. How would you characterize teacher views regarding the inclusion of class discussion in student grades in your past college courses? (.657)	2.77	3.15	3.08 [†]
36. Which of the following best express your view of the long-term value of learning to express one's views in public? (.455)	3.67	3.86	4.10 [†]
42. How would frequent participation in college classes likely affect your grades in those courses? (.651)	3.54	4.00	4.25 ^{††}

Factor 2 Total	24.21	28.19	30.13
----------------	-------	-------	-------

Items on Factor 3-	Low	Medium	High
---------------------------	------------	---------------	-------------

Personal Benefits of Participation:

21. What effect does frequent discussion by other students have on your concentration in class? (.739)	3.38	3.34	3.68
--	------	------	------

22. What effect does your personal participation in class discussion have on your concentration in class? (.620)	3.31	4.07	4.32 ^{††}
--	------	------	--------------------

23. What effect does the option of volunteering comments whenever you wish have on your concentration in class? (.686)	3.60	3.57	3.77
--	------	------	------

24. How does the possibility that you might be called on to respond to an instructor question affect your concentration in class? (.452)	3.50	3.73	3.89
--	------	------	------

25. How does frequent discussion in the class as a whole affect your enjoyment of a class? (.608)	3.21	3.81	4.19 ^{††}
---	------	------	--------------------

29. To what degree does student sharing of personal experiences in class discussion contribute to the quality of the discussion? (.725)	3.54	3.74	3.91
---	------	------	------

44. How much of a personal priority is improving the amount and/or quality of your participation in class discussion? (.646)	3.19	3.53	3.60 [†]
--	------	------	-------------------

Factor 3 Total	23.73	25.79	27.36
----------------	-------	-------	-------

All other items:

	Low	Medium	High
5. Which of the following best describes your perspective of the relevance of your comments in class discussion?	4.04	4.24	4.25
6. Which of the following most accurately describes your typical level of preparation for class discussion in past courses?	3.90	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?	3.85	4.32	4.36
10. How do you feel when an instructor indicates at the beginning of a course that class discussion is welcomed but optional?	4.06	4.15	3.89
11. What would be the relative likelihood of your asking a question versus answering a question in class discussion?	3.31	3.47	3.02
13. How do you think that your keeping a record of your participation in class discussion would affect your concentration on the discussion?	2.96	3.25	3.06
15. What do you see as the relationship between the quantity and quality of class discussion?	2.79	3.02	2.92
17. How do you feel about a discussion format in which students volunteer comments rather than being called on by the instructor?	4.06	4.05	4.15
18. Who is responsible for a high level of student participation in class discussion?	3.42	3.64	3.49

19. Who is responsible for very limited student participation in class discussion?	3.19	3.51	3.36
26. How do you feel toward students who frequently comment in class discussion?	3.12	2.88	3.45
27. What are your academic expectations of students who frequently participate in class?	4.25	4.32	4.36
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?	4.40	4.53	4.62
30. How do you typically respond when no one else is responding to a teacher question?	2.98	3.51	4.28
32. What effect do long pauses between teacher questions and student responses have on your desire to participate in class discussion?	2.65	3.07	3.38
35. In comparison to other classes you are taking this semester, what expectation do you have for your participation in 210 class discussion?	3.56	3.73	3.92
37. Your interpretation of how teachers feel about class discussion is best reflected in which of the following claims?	4.21	4.36	4.17
39. Many teachers try to stimulate class discussion by asking questions. Which of the following best expresses your view of most teacher questions?	3.42	3.46	3.68

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?	3.17	3.53	3.51
41. Which of the following best represents your ability to judge the relevance of your comments in class discussion?	4.19	4.42	4.36
43. How do you typically feel when you have volunteered a comment in class discussion?	3.19	3.58	3.64
45. What is your opinion of the social status of students who participate frequently in class discussion?	3.21	3.15	3.30
46. What is your opinion of the social status of students who participate little, if at all, in class discussion?	2.85	2.78	2.72
47. What is your view of the relationship between the perceived relevance of course content and student inclination to participate in class discussion?	3.83	4.44	4.49
48. Which of the following best represents how students' participation in class discussion will affect their personal standing with teachers?	4.00	4.42	4.25
49. Which of the following best represents the effect of a teacher's friendliness on student participation in class discussion?	4.04	4.17	4.15
50. Compare the effects of teacher friendliness and teacher knowledge of the subject matter in the course on student participation in class discussion.	3.23	3.20	3.25

All Discarded Items Total	95.88	100.97	101.65
---------------------------	-------	--------	--------

[†]Low and High groups are significantly different ($p < .05$). Note that for item 34, only the Low and Medium groups are significantly different ($p < .05$).

^{††}Low and High, and Low and Medium groups are significantly different ($p < .05$).

^{†††}Low and High, Low and Medium, Medium and High groups are significantly different ($p < .05$).

Vita

Katherine Rae Krohn was raised in Springfield, Ohio and graduated from Catholic Central High School in 2001. She received a B.A. in Psychology with a minor in Child Studies from Miami University in Oxford, Ohio in 2005. She enrolled immediately at the University of Tennessee to pursue her Ph.D. in School Psychology. In December 2008 she received a M.S. in Educational Psychology. Katherine completed requirements for her Ph.D. in August 2010 after a year-long internship accredited by the American Psychological Association.