The Dissertation Committee for Jason Robert Crandall certifies that this is the approved version of the following dissertation:

FLIPPING AN LSAT COURSE: A DESIGN STUDY INVESTIGATION OF COLLABORATIVE INSTRUCTION AND INVERTED CURRICULUM IN A TEST PREPARATION COURSE

Committee:

Diane L. Schallert, Supervisor

Edmund T. Emmer

Erika A. Patall

Marilla D. Svinicki

Gary Keith

FLIPPING AN LSAT COURSE: A DESIGN STUDY INVESTIGATION OF COLLABORATIVE INSTRUCTION AND INVERTED CURRICULUM IN A TEST PREPARATION COURSE

by

Jason Robert Crandall, B.A.; M. A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2014

Acknowledgements

First and foremost, I must thank my wife, Mary Ann. We began this journey before our marriage and ended it with two wonderful children. Your love, your support, your encouragement, and at times, your stubbornness made this possible. I can't imagine what it would have been to try to do this without you and I'm glad that I don't have to. I love you.

To Rowyn and Avery, I'm sorry that I missed out on some of your bedtimes and couldn't always stop working to play with you. But now I'm that I'm done with my big paper, I'll do my best to make it up to you.

To Erika, Ed, and Gary, thank you for your willingness to serve on my behalf. I value your input and guidance, and look forward to your feedback on my work here. In particular, I wish to thank Gary for finding time to share your research with me and notes of interest on pre-law students and flipped instruction, even when I was often slow to reciprocate. Your thoughts did not go unnoticed or unread.

Special thanks to Marilla for involving me in the HIT project. That experience has been invaluable to me during my job search process and has impressed all of my interviewers. Thank you for recommending me and for showing me what expertise in instructional design and curriculum planning looks like.

Diane, thank you for your support and your encouragement. I always left our meetings with not only a clearer sense of the work I was expected to do here, but also an extra measure of motivation. When my enthusiasm waned, yours never flagged and for that, I am deeply indebted. For all that I have learned from you, I feel some remorse that I never bothered to ask you to explain to me the difference between *while*, *although*, and *whereas*.

Thank you to Dan Robinson, Camilo Guerrero, Sarah Jones, Alyssa Reinhart, and Daniel Clark. I couldn't have asked for a better group of collaborators to help me understand

how to ask better questions and seek better answers. Your example and your friendship have meant a great deal to me and I look forward to future collaborations.

Jane Vogler Cragun, thank you for unwittingly allowing me to use your dissertation as a guide for my own. I'm glad I picked a good role model.

Linda Dickens, I couldn't have asked for a better, more supportive boss during the past year. Your flexibility and encouragement made this process immeasurably easier than it would have been otherwise.

Rob Donald, your optimism and professionalism made every interaction a pleasure and your technical expertise was greatly appreciated. Thank you.

I am also indebted to Virginia Stockwell, whose prompt and knowledgeable guidance greatly enhanced my graduate school experience. I can't fully understand how you get quite so much done so quickly, but I'm glad to have benefitted from it.

Finally, I must thank my parents, Alan and Sandy Crandall, for inspiring me to be curious about the world and strive for greatness. Thank you for raising me to love learning and to accept challenges.

Flipping an LSAT course: A design study investigation of collaborative instruction and inverted curriculum in a test preparation course

Jason Robert Crandall, Ph.D.

The University of Texas at Austin, 2014

Supervisor: Diane L. Schallert

A primary goal of education is not only to inform but to transform learners. As instructors shift their focus from a one-size-fits-all emphasis on content delivery to a flexible, student-centered approach, questions of student engagement and student motivation become key. In many educational settings, instructors are faced with a classroom of students with varying, and often unknown, levels of motivation, ability, and commitment. Effectively addressing the educational needs of such a range of students often requires significant changes to traditional pedagogy.

A recent pedagogical design that has been facilitated by the advent of easily accessible and low-cost multimedia technology is the "flipped classroom," a course structure that asks students to view lectures prior to class and replaces the traditional in-class lecture with collaborative, problem-based instruction. The aim of the present study was to explore the experience of introducing a flipped curriculum into a LSAT (the nationally used entrance exam for admission to law school) preparation classroom. The study used a design research approach to investigate two iterations of the flipped

curriculum across three courses.

Quantitative and qualitative data were used to describe the experience of a flipped curriculum for both the instructor and the students. When compared to a traditional curriculum, results showed no significant effect on overall test score improvement, but students in the flipped courses did show greater improvement than those in a traditional course on one of the three LSAT section subscores. The results also showed that students in flipped courses had marginally lower overall attendance, greater classroom community, high levels of engagement, and moderately high belief in group effectiveness.

Table of Contents

List of Tables

List of Figures

Chapter One: Introduction

Rationale

A primary goal of education is not only to inform but also to transform learners. Dedicated instructors strive to help students learn content, as well as develop effective ways of learning and engaging content. As instructors shift their focus from a one-size-fits-all emphasis on content delivery to a flexible, student-centered approach, questions of student engagement and student motivation become key. In many educational settings, instructors are faced with a classroom of students with varying, and often unknown, levels of motivation, ability, and commitment. Effectively addressing the educational needs of such a range of students often requires significant changes to traditional pedagogy.

One of the more robust pedagogical adaptations in recent history has been the use of collaborative learning. Indeed, the research and application of collaborative learning has been described as "one of the greatest success stories in the history of educational research" (Slavin, 1996, p. 43). Gilles and Ashman (2003) defined collaborative learning as "the instructional use of small groups so that students work together to maximize their own and each other's learning" (p. 168). Collaborative learning has a broad and widely-supported base of research in various disciplines and settings, including the STEM (science, technology, engineering, and mathematics) fields (Springer, Stanne, & Donovan, 1999), economics (Johnston, James, Lye, & McDonald, 2000), lifelong learning (McConnell, 1999), teacher development (Bray, 2002), and second language (L2) learning (Oxford, 1997). In a study regarding how to support peer learning through

proper assessment, Boud, Cohen and Sampson (1999) noted that the key educational outcomes of peer learning include first and foremost "the development of learning outcomes related to collaboration, teamwork, and becoming a member of a learning community" (p. 415).

In a meta-analysis of collaborative learning studies in science, technology, engineering, and mathematics classrooms, Springer, Stanne, and Donovan (1999) stated that the success of small group learning depends upon the learner's association of the group structure with successful learning. They described three important components necessary for increasing the probability that small group learning will be successful: Students must perceive the effect of group learning on (1) cognitive elaboration, (2) affective collaboration, and (3) motivation. Cognitive elaboration involves opportunities for students to think about and engage with classroom content. Collaborative learning can promote cognitive elaboration by increasing the amount of time students spend on task, encouraging greater interaction with the instructor, and fostering peer-to-peer interaction. As a result of such interactions, "cognitive conflicts will arise, inadequate reasoning will be exposed, and enriched understanding will emerge" (Springer, Stanne, & Donovan, p. 25). Affective collaboration can also result from collaborative learning as students may feel less pressure when performing in front of the small group as compared to the entire class. Also, collaborative learning in small groups may increase opportunities for members of underrepresented groups to express opinions and contribute to the classroom learning atmosphere (Springer, Stanne, & Donovan, 1999). Finally, motivation can be enhanced or diminished by the effectiveness of the collaborative learning structure

selected by the instructor, as discussed below.

In a standard university classroom, successful implementation of collaborative learning is often linked to the grading structure. Poor use of grading and assessment with collaborative learning can encourage free riding, unequal distribution of effort, decreased support of individual and group learning goals, and general antipathy toward group learning (Feichtner & Davis, 1984-85). Careful consideration of incentive and reward structures can ameliorate any or all of these problems, and instructors are advised to review research on small group grading before adopting this strategy into their courses. According to Johnson and Johnson (1998), when appropriate classroom structures are used,

> there is considerable evidence that students will exert more effort to achieve (learn more, use higher-level reasoning strategies more frequently, build more complete and complex conceptual structures, and retain information learned more accurately), build more positive and supportive relationships (including relationships with diverse individuals), and develop in more healthy ways (psychological health, self-esteem, ability to manage stress and adversity). (p. 73)

However, not all instructors have recourse to grades as an instructional tool. For example, instructors who teach courses to adult learners typically do not grade their students. In some lifelong educational settings, successful course completion may result in certification or a similar credential from the instructor, but in many others, the motivation for and ultimate goal of the instruction is extrinsic to the course itself. Bar review courses, certified public accountant exam preparation, real estate continuing education courses, project manager certification, and nursing exam review courses are all examples of adult education settings that are not typically graded and where success for the learner is defined primarily or exclusively by achievement on a non-course activity, such as a

high-stakes exam external to the course.

Another common feature of such courses is a tightly prescribed curriculum. Because the ultimate learning objective is external to the course, instructors in these educational settings have limited autonomy over what material to include and what to leave out. The pressure to cover "everything" can be amplified in a commercial setting, where students have paid for a specific course and may expect the instructor to deliver a steady flow of information and expertise. Finding time to create an active, collaborative learning environment that emphasizes problem-solving and deep processing while covering the required material and satisfying the students' expectations can be very challenging.

One potential solution to this challenge is the use of a *flipped classroom*, which inverts the lecture-homework cycles found in most traditional classrooms (Lage, Platt, & Treglia, 2000; Strayer, 2007). Rather than introducing content to students in the classroom through lecture and asking them to engage with it more deeply out of the classroom through homework, a course using a flipped curriculum introduces content out of the classroom through recorded lectures and asks students to engage the content more deeply in the classroom through problem-based activities. Such a switch "impacts the roles and responsibilities of instructors and students alike and holds the potential for improving classroom-based learning experiences for both groups" (Brown, 2012, p. 8). Of particular note is that this model can be used to increase the exposure of students to relevant problems within a specific educational domain without reducing or materially altering the delivery of expected information (Baker, 2000).

However, according to a review of the literature, there are few peer-reviewed publications of any kind on flipped curriculum designs and even fewer published, controlled experiments comparing course instruction in a flipped curriculum to a traditional, lecture-based curriculum (Bishop & Verleger, 2013). My own search indicated dozens of blogs, magazine articles, and conference presentations on the use of a flipped curriculum among university and high school students, but no studies involving adult learners. Despite increasing popularity and even recommendations for best practices in implementing this design, research on the effects of a flipped curriculum in authentic learning environments is scarce, at best.

This study investigated the efficacy of a flipped curriculum and collaborative learning design in an ungraded, high-stakes adult learning course. Specifically, I introduced a flipped curriculum into a commercial LSAT (Law School Admissions Test) prep course. Three sections of the course replaced classroom lecture with small-group collaborative learning and were compared to traditional course instruction in terms of the effect on student engagement and student learning. In order to facilitate meaningful group work and ensure that students receive the full curriculum as intended by the test preparation company, I prepared lecture videos for students to watch prior to class to make room for problem-based instruction during class. Finally, due to resource considerations and constraints on available control group data, I used a design experiment process to adapt the instruction to students' needs and to document all changes made to the implementation during the course of the study.

Educational Setting

Standardized test preparation is a somewhat unusual educational setting, as students in these courses are motivated not by a desire to succeed in the course itself or to earn the instructor's approbation, but by their self-determined perceived need to earn a higher score on an extra-curricular test. There is very little research on learning process in an instructional test preparation course. Although ETS and others have conducted extensive research on the validity, reliability, interpretation, and use of standardized test results (e.g., https://www.ets.org/gre/research; http://www.lsac.org/ lsacresources/research/all; http://research.collegeboard.org/ programs/sat/data/validity-studies), there is also an extensive line of research describing the negative effects of standard-based legislation and testing on students, teachers, and administrators (Ravitch, 2010; Haladyna, Haas, & Allison, 1998; Randall, 2006). Other researchers have commented on the questionable ethics of teachers coaching students on high-stakes tests in K-12 settings (Smith, 1991). Researchers, administrators, and teachers largely agree that an increased emphasis on preparing primary and secondary students for high-stakes tests hinders learning and interferes with educational progress (McNeil, 2000).

However, post-secondary test preparation, such as that for MCAT, GRE, or LSAT exams or professional licensing exams, differs from teaching to the test in K-12 settings in several significant aspects. First, post-secondary test preparation is typically conducted outside of school and seldom by instructors who grade their students. Second, unlike many state educational agencies, the test-development institutions that create post-secondary standardized tests typically offer previously-used questions to students and

instructors on a licensed basis. Commercial test prep companies can purchase many or most of the previously administered items, rather than simulating or stealing them. Finally, post-secondary test prep is purely voluntary and does not preempt other forms of instruction. As in other informal education settings (such as professional certifications), test preparation students can select their own methods of preparation and typically have narrow educational objectives.

Notwithstanding these differences, very little research exists on post-secondary test prep courses. Much of the extant research argues that expensive, time-consuming courses are of debatable value relative to self-study (Berkowitz, 1998). However, some recent neurological research found that not only can intensive prep courses significantly increase scores, but they may also result in detectable changes to brain structures (Mackey, Whitaker, & Bunge, 2012). In my literature search, I was unable to find any research regarding the effectiveness of any instructional strategies within these courses.

Despite the sparse and apparently contradictory findings regarding the efficacy of these courses, their use remains very widespread. In 2011, the Law School Admissions Council (LSAC, providers of the LSAT) commissioned a study of self-reported method of test preparation for LSAT takers from 2008-2010 (Evans, Sweeney, & Reese, October 2011). Approximately 35% of the 201,748 students questioned indicated that they had used a commercial test-preparation service. The existence of this large and relatively understudied population of students presents an opportunity to investigate the applicability of well-understood educational interventions in a non-school setting.

The Study

In order to facilitate group work within the classroom, I modified the existing course structure of the LSAT prep course of which I was the official instructor. As traditionally constituted, the course consists of 12 four-hour lessons, each devoted to two or three common question types found on the LSAT. Lessons are typically divided into two main portions, each of which is focused on one of the three section types. For example, the first half of lesson 1 covered inference questions from the logical reasoning section, and the second half covered sequence games from the analytical reasoning section. The full course syllabus can be found in appendix A. Each portion begins with a lecture regarding the features of and recommended approaches for a given question type. Following the lecture, the instructor models the approach using practice questions and allots the remainder of the lesson to individual, timed practice questions, followed by lecture-style explanations for each question. Between lecture, modeled instruction, and individual practice, the traditional course structure leaves no time for collaborative work. Thus, adoption of a flipped curriculum was seen as allowing room during class meeting for collaborative activities during class without altering the content covered or sequence of instruction in the classroom.

This study compared results from past courses, taught in the traditional lecture-based method, to three courses featuring small group collaboration and a flipped curriculum style. For each lesson, students were asked to read the lesson text prior to class and watch a brief (15 to 20 minutes) video containing supplemental material that I, as the instructor, traditionally delivered during class. Students were asked not to attempt

problem sets from the lesson and those who wished to practice techniques from the video and the written lecture prior to class were prompted to use homework problems. Classes typically began with a readiness assurance quiz to check reading, attendance, and general comprehension. Questions for this quiz were primarily recall and recognition based, and the results determined what topics, if any, I reviewed with the entire class before proceeding to modeled instruction and small group work. Once I determined that students had sufficient background knowledge to attempt practice questions, the students were divided into small groups (typically three to four members per group). Specific group instructions varied depending on question type, but all classroom activities were focused on encouraging collaboration and deeper processing of critical reasoning concepts.

Student responses for all questions were gathered using a web-based classroom response system (specifically, I used Socrative, which is a free, cross-platform system that allows synchronous data collection for up to 50 students on any web-enabled device). In addition, I administered several surveys to students in these courses: the first survey was given to students immediately after the initial practice LSAT and served primarily to gather data related to demographic and learning goals. The second survey was administered after the sixth lesson and asked students for their opinions about course structure and teamwork. A final survey was given after the courses ended and asked questions about course engagement. These results were compared to results from students in previous traditional courses that I taught. Finally, to provide a detailed view of students' thinking during the course, I recorded and coded group discussion during three of the lessons.

Research Questions

Question 1: Compared to traditional LSAT prep instruction, what effect does a flipped curriculum structure have on student learning, student engagement, and student retention?

Question 2: What is the experience, for the instructor and students, of collaboratively learning in an LSAT prep course?

Question 3: What does group discourse reveal about the learning environment and process within a course using a flipped curriculum?

Research Perspective

My interest in pre-law students stems from teaching LSAT prep courses and offering private tutoring for the past eight years to hundreds of students. During these courses, more than 30% of students failed to complete all course requirements. I have observed that whereas many low-scoring students persist in coursework and test preparation, a few high-scoring students abruptly stop attending courses. However, there is a high correlation between initial test score and course persistence (among my students, regressing initial test scores onto total number of in-class practice tests taken has an R^2 of 0.952). Many students have noted that the pace of the course is often too fast for them and that they feel overwhelmed by the amount of material covered. Others have questioned why the class does not always cover all of the material in each lesson. I hypothesized that flipped instruction and small group collaborative problem-solving could address both concerns, by providing additional time-on-task for problems during class, while adjusting the pace to the particular needs of each class member.

This research was meant, in part, to test Slavin's (1996) theories regarding group goals and individual accountability as prerequisites for successful collaborative instruction. If a course designed to include collaborative group work proved to be successful in this setting, it would suggest that group learning can be effective even when there is no overt mechanism for encouraging group accountability and even among students, adult learners, with many years of successful formal instruction and with explicitly individualistic learning goals. It was also meant to test the efficacy of the flipped curriculum design among self-directed adult learners and to add to the narrow body of research on this instructional approach. My hope was that the results of this study would offer guidance to instructors and curriculum designers who wish to provide active, collaborative environments to adult students in self-selected learning environments without sacrificing the content they receive.

Chapter Two: Literature Review

This review is intended to provide a basic overview of the theoretical developments and empirical research relevant to explaining pedagogical approaches and applications of collaborative learning, problem-based learning, and learning with a flipped curriculum. Specifically, I discuss how the theory of socio-constructivism has informed the development of active, student-centered classroom environments, such as those intended to result from collaborative learning course designs. The review also presents some of the research on adult learning environments, specifically pre-law education and test-prep education. Finally, I review studies on student engagement and on student beliefs regarding team effectiveness in order to explain the outcome measures chosen for this study.

Theoretical Background

Much of the modern thinking on student-centered classrooms and socio-constructivism can be traced back to the writing of Lev Vygotsky, a Russian psychologist whose writings from 1925 to 1934 continue to influence modern thinking about education. Vygotsky was concerned with the relationship between speech and thought and the social nature of learning and wrote extensively on these topics. As Vygotsky's work was translated and introduced to Western academia during the 1970s, it began to influence classroom teaching, curricular theory, and developmental psychology.

Two of Vygotsky's concepts that are most relevant to the topics of student engagement and active learning are the Zone of Proximal Development (ZPD) and scaffolding (Vygotsky, 1978). The ZPD is defined as a learner's developmental level in a

given learning domain such that the learner can learn from and accomplish certain tasks only with the measured assistance of someone who is more proficient in that domain. Vygotsky maintained that the majority of instruction should take place within this zone as the student will eventually learn to do these tasks alone and continually expand his or her ZPD into progressively more challenging cognitive domains. Furthermore, tasks within the ZPD are neither so difficult as to be demoralizing nor so easy as to be uninspiring.

A closely related concept is the process of scaffolding (Vygotsky, 1978). Scaffolding is the process by which more experienced individuals modify tasks that are too difficult for learners to complete without their assistance. Such modifications can include completing the task together, simplifying the task, carefully guiding the student through each step of the task, or completing challenging portions of the task in advance before the student begins. In each case, the degree of scaffolding required is inversely proportional to the student's task ability; ultimately, the fully competent student should complete the task with no scaffolding. Tasks for which the student requires some degree of scaffolding are considered to be within the student's ZPD. Referencing Brophy (1999) and Pintrich and Schunk (2002), Pintrich (2003) wrote:

> Many motivational theories, as well as cognitive theories (including Vygotskian models), stress the importance of providing tasks that are within the range of competence for students. The tasks should be neither too easy nor too difficult, but challenge students in appropriate ways. (p. 672)

As mentioned above, Vygotsky's writings were highly influential in the development of a view of learning referred to as socio-constructivism. Socio-constructivists believe that knowledge about the world is constructed through shared cultural experiences and that the meaning of learning artifacts can only be established through interactions among

teachers, learners, and subjects. One of the most important methods of such interaction is classroom discussion. Discussion is the simplest way to establish how much students understand, to create new knowledge about learning artifacts, to diagnose a learner's ZPD, and to determine the appropriate amount of instructional scaffolding. By facilitating more and better classroom discussion, well-planned group instruction can allow an instructor to identify more easily a student's ZPD and to offer individualized scaffolding.

Of course, most classrooms are far from ideal. Nystrand's (1997) landmark book, Opening Dialogue: Understanding the Dynamics of Language and Learning in the English classroom, investigated classroom discussion during a three-year period in over 60 classrooms with more than 2400 students. Nystrand found that instructors averaged three minutes of classroom discussion per 60 minutes of instruction. McKeachie and Svinicki (2006) discussed various barriers to classroom discussion, including "students' feeling that they are not learning," "the instructor's tendency to tell a student the answer" too quickly, instructor discomfort, and the difficulty of "appraising the group's progress," and not "be[ing] aware of barriers . . . that are blocking learning" (pp. 44-45). Allowing true discussions is particularly challenging in adult learning environments, where instructors may have little means to motivate classroom participation and some students may altogether reject the instructor's attempts to promote active learning.

Active Learning

Most research in active learning is informed, to some degree, by Chickering and Gamson's (1987) Seven principles for good practice in undergraduate education. These principles particularly emphasize the frequency and quality of student-faculty contact and

student-student contact, and are relevant for adult learners as well as undergraduate students. As Guerrero (2009) wrote, "Each of these principles rests on the belief that students benefit from an instructor's ability to design a learning environment that considers the students' activity level, cooperation, diversity, expectations, interactions, and personal responsibility for learning" (p. 7). According to Chickering and Gamson (1987), good teaching practice:

1. Encourages student-faculty contact
2. Encourages cooperation among students
3. Encourages active learning
4. Gives prompt feedback
5. Emphasizes time on task
6. Communicates high expectations
7. Respects diverse talents and ways of learning (pp. 1-2)

Each of these points warrants brief consideration in terms of their individual impact on creating a positive learning atmosphere and a student-centered classroom.

Student-faculty contact is described as the "most important factor in student motivation and involvement" (Chickering & Gamson, 1987, p. 1). Students who have frequent and diverse contact with faculty tend to be more involved and more successful than other students (Edmonds & Edmonds, 2008; McKeachie & Svinicki, 2006). One of the difficulties in a traditional classroom is the limited degree of student-faculty contact. Nearly all contact is unidirectional (from teacher to student) and is seldom personalized. Office hours, effectively moderated discussions, lab sessions, and interactions immediately before or after class can help address this difficulty, but repurposing the classroom altogether by "flipping" the curriculum and using class time strictly for group-

based problem-solving could significantly increase the quality and frequency of student-instructor contact.

In addition to contact with faculty, contact with other students can be highly beneficial to students. Chickering and Gamson wrote, "Good learning, like good work, is collaborative and social, not competitive and isolated" (1987, p. 1). Thus, effective undergraduate learning environments should seek to *encourage cooperation among students*. Various instructional methods, such as team-based learning (Michaelsen & Black, 1994; Michaelsen, Knight, & Fink, 2002), depend heavily on the learning benefits derived from small-group interactions. Even a less formal approach, such as asking students to discuss questions with their neighbors prior to initiating class-wide discussion, allows students to create knowledge together and receive low-stakes feedback from their peers.

Active learning has been defined as "instructional activities involving students in doing things and thinking about what they are doing" (Bonwell & Eison, 1991, p. 1). Bonwell and Eison (1991) added: "To be actively involved, students must engage in such higher-order thinking tasks as analysis, synthesis, and evaluation" (p. 1). Chickering and Gamson (1987) claimed that "students do not learn much just sitting in classes listening to teachers. . . . They must make what they learn part of themselves" (p. 1). Modifying traditional classroom environments to allow active learning is central to the effective application of collaborative pedagogy and will be discussed further in the "Small Group Learning" section below.

Another core principle of student-centered learning is *giving prompt feedback*.

Chickering and Gamson (1987) wrote: “Knowing what you do and don't know focuses learning. . . . In classes, students need frequent opportunities to perform and receive suggestions for improvement” (p. 1). Such regular feedback helps both students and instructors actively to adapt the learning environment. Students who discover that they have sufficiently understood lecture material may be motivated to continue engaging in the course, whereas students who receive feedback that their understanding is incomplete may be motivated to seek additional contact with their peers or with the instructor. The design used in this study emphasized prompt and frequent feedback to students.

Emphasizing time on task and *communicating high expectations* are linked to each other in that many instructors expect students to remain on task throughout course instruction. Chickering and Gamson (1987) argued that “how an institution defines time expectations for students, faculty, administrators, and other professional staff can establish the basis for high performance for all” and that “expecting students to perform well becomes a self-fulfilling prophecy when teachers and institutions hold high expectations of themselves” (p. 2). Using small-group instruction along with a flipped curriculum design allows students to spend more time on task than in a lecture classroom and also allows instructors to set higher standards for the learning process collectively than might be possible individually, particularly when student ability levels vary significantly.

Finally, student-centered learning environments should *respect diverse talents and ways of learning*. The good practices described above “work for many different kinds of students--white, black, Hispanic, Asian, rich, poor, older, younger, male, female, well-

prepared, under prepared. But the ways different institutions implement good practice depends very much on their students and their circumstances" (Chickering & Gamson, 1987, p. 3). What constitutes active learning for one student or one population of learners may not always be reflected by other students. Small-group learning techniques, such as those used in this study, have been shown to expose students to viewpoints that may differ substantially from their own and that may supplement the instructor's perspective.

Small Group Learning

For the past four decades, researchers have investigated the use of small groups in various learning environments as a means to encourage active learning and to apply principles consonant with socio-constructivist views of learning. Small group instruction has been found to benefit learners across content domains and at all levels of education (Barron, 2000; Colbeck, Campbell, & Bjorklund, 2000; McConnell, 1999). Various features of group structure and process have also been examined, such as group composition (Webb, Nemer, Chizhik, & Sugrue, 1998), group roles and responsibilities (Saleh, Lazonder, & de Jong, 2007), group discourse (Barron, 2003), and group size (Fuchs, et al., 2000). Over time, group work research has also focused on specific, formal implementations such as collaborative learning, peer instruction, and team-based learning, each of which will be addressed in the following sections.

Collaborative learning. Among the influential efforts to describe and formalize the implementation of group learning in classroom settings is the work of Slavin (1991; 1996). Slavin used the term cooperative learning to describe the use of small groups of students learning together and teaching one another. Although his research focused

primarily on elementary and secondary students, the principles he recommended are applicable to learners at all levels. Slavin emphasized that successful cooperative learning requires more than simply requiring students to discuss material with each other or to turn in joint assignments. Instead, he argued, "Two elements must be present if cooperative learning is to be effective: group goals and individual accountability...groups must be working to achieve some goal or to earn rewards or recognition, and the success of the group must depend on the individual learning of every group member" (1991, p. 76). Slavin reported that when such structures were in place, cooperative learning had positive effects on learning and on "such diverse outcomes as self-esteem, intergroup relations, acceptance of academically handicapped students, attitudes toward school, and ability to work cooperatively" (p. 71).

Collaborative learning was also noted by Chickering and Gamson (1987) as a possible approach for encouraging active learning in the undergraduate classroom. A broad body of research into collaborative learning among undergraduate students has found positive student outcomes (Astin, 1993; Kuh & Vesper, 1997), although not without caveat. Astin (1993), in particular, noted that although collaborative learning was associated with increases in student-student contact and student-faculty contact, some faculty efforts to promote active learning were negatively associated with student retention. One possible explanation for this is the notion presented by Johnson, Johnson, and Smith (1991) that poorly designed or poorly implemented group work may be worse than traditional, lecture-based instruction. Accordingly, current literature on collaborative learning typically includes very specific recommendations for proper implementation and

course design.

Peer Instruction. Another form of small group learning that informed the course design used in this study is peer instruction (Mazur, 1997). Peer instruction (PI) is a socio-constructivist approach that depends upon students to teach each other core course concepts. During PI, lectures are divided into a series of short presentations, followed by a conceptual question, termed a *ConcepTest* by Mazur (Crouch & Mazur, 2001). Typically, students are asked to consider the question, formulate an answer, and report their answer to the instructor, often using clickers or some other classroom response system. The instructor may choose to reveal these results to the class, and then always asks students to discuss their choices with each other. Crouch and Mazur (2001) wrote that "the instructor urges students to try to convince each other of the correctness of their own answer by explaining the underlying reasoning" (p. 970). After a few minutes of peer instruction, the students report revised responses to the instructor, who offers any final comments or explanation before moving to the next presentation.

In a review, Crouch and Mazur (2001) reported results from ten years' worth of peer instruction in calculus and algebra-based introductory physics courses. Across a variety of measures, including two standard tests and traditional questions, students taught through peer instruction showed higher normalized gain scores than those taught in a traditional course. These effects were independent of instructor and were present for several different variations of the peer instruction process (p. 970). Peer instruction is especially well-suited to resolving widely-held misconceptions, as students may better understand why their peers hold incorrect beliefs than an instructors does. My study used

peer instruction techniques when appropriate to address a specific question.

Team-Based Learning. Another specific group learning strategy is team-based learning (TBL), developed by Larry Michaelsen (Michaelsen & Fink, Preface, 2008). In order to be fully implement, TBL requires a) strategically formed, permanent teams, b) readiness assurance, c) carefully selected application activities, and d) peer evaluation (Michaelsen & Sweet, 2011). Michaelsen and Sweet argued that TBL should not be seen as a variation of cooperative learning that can be "dropped into" a lecture-based course. Rather, TBL "is a comprehensive instructional system that, when implemented correctly, achieves an increasingly interlocking synergy and amplifies students' social and intellectual capacities over time" (p. 43). They also reported research findings to the effect that TBL had positive effects on "test performance, engagement, retention, student attitudes toward group work, and student satisfaction with their learning experience." (p. 50). My study used readiness assurance principles for both semi-permanent groups and variable groups solving challenging application activities. However, the groups were not strategically chosen and students were only asked to evaluate their group processes, rather than their individual peers. Thus, the design likely lacked some of the effectiveness of a more fully integrated TBL approach.

Research on group differences. One important study that helped guide this study's analysis was Brigid Barron's investigation of differences in problem-solving ability among groups of gifted learners (2003). Barron argued that "there is need for a better articulation of the characteristics of interactions that lead to differentially productive joint efforts" (p. 309). Her analysis focused on a sample of sixth-grade triads

given a challenging math program, and utilized measurement constructs that focused on group interaction. With respect to why different groups may succeed or fail with similar prior knowledge and the same task, Barron offered three main observations: First, joint management of attention is a fundamental, and often challenging, aspect of collaborative problem solving. Her research suggested that the "achievement of joint attention was consequential from problem solving and learning" (p. 310). Second, speakers and listeners alike share the responsibility of creating and maintaining joint attention. Some participants struggled to have their ideas heard, particularly if their partners were self-focused, whereas others were able to make themselves heard through resolute effort. Third, issues of identity, self, and relational context must be addressed to explain group differences effectively. Barron wrote that coordinated mutual engagement is "both an interpersonal- and content-related process. The less successful groups in this data set exhibited relational issues that challenged mutual engagement and prevent the group from capitalizing on the insights of members" (p. 311). Each of these themes was observed in some form among the groups in this study.

Flipped Curriculum

The concept of a "flipped" curriculum is based on the idea that class time may be better spent by students actively working on problems during class rather than listening to a lecture. In its simplest form, a *flipped curriculum* means that "events that have traditionally taken place inside the classroom now take place outside the classroom and vice versa" (Lage, Platt, & Treglia, 2000, p. 12). Many early applications of flipped curriculum designs were in the field of engineering. In order to satisfy accreditation

guidelines from the Accreditation Board for Engineering and Technology (ABET), some instructors adopted problem-based learning methods in their classrooms. However, the nature of the curriculum requires a significant transfer of information from instructor to student and forced these instructors to consider alternative methods of delivering lecture material to their students. Pipes and Wilson (1996) were among the early promoters of recording video lectures for students to view prior to class and devoting in-class instruction to small group-based problem-solving. The rationale of ensuring that students would be able to cover the same material as in a lecture classroom was also my primary motivation for adopting this course design in my study.

In recent years, the concept of a "flipped curriculum" has become popular in secondary and post-secondary settings. For example, the Center for Teaching and Learning at the University of Texas at Austin has investigated the process of flipping classrooms on campus and has offered support to faculty who wish to attempt this course design. There are also numerous magazine articles, blogs, and books designed to help instructors "flip" their classrooms. One main driver of the current interest in flipped curriculum designs is the increasing ubiquity of educational technology that allows broader access to course content and increased customization in course delivery (Woolf, 2010). Tools such as web-hosted video lectures, interactive course modules, adaptive tutoring systems, and course management systems are central to modern implementations of the flipped curriculum. However, according to a review of the literature, there are few peer-reviewed publications of any kind on flipped curriculum design (Bishop & Verleger, 2013) and even fewer published, controlled experiments comparing course instruction in

a flipped course to a traditional course.

One example of such a controlled experiment is Strayer's (2012) study. Strayer studied students in two sections of his introductory statistics courses, one traditional with in-class lectures followed by homework problems, and one flipped with homework lectures preceding in-class problems. The students in the flipped condition learned the lecture material through an intelligent tutoring program. Strayer wrote that, compared to the students in the traditional course, "students in the inverted classroom were less satisfied with how the classroom structure oriented them to the learning tasks in the course, but they became more open to cooperative learning and innovative teaching methods" (p. 140). He suggested that integration of the face-to-face and online portions of a flipped course is essential in helping students adjust to an unfamiliar use of classroom time. He also noted that an introductory course may not be the ideal setting for a flipped course, as novice students may find the time and effort needed to connect the various portions of the course frustrating without a vested interest in the material.

Another such study was conducted by Davies, Dean, and Ball (2013) who compared a traditional condition, a flipped condition, and a simulation-based condition for teaching an introductory course in using Excel. The authors were particularly interested in the efficiency and scalability of the flipped and simulated conditions. In the flipped condition, students were given video lectures and supplements, along with optional remedial lectures and classroom support, which required fewer computers and software licenses than a traditional course. The simulated classroom was delivered entirely online, with students using a web-based spreadsheet simulator to complete pre-

determined objectives and tasks. The simulation software provided videos detailing the use of menus and functions to perform the tasks, but these did not receive video lectures.

Davies, Dean, and Ball (2013) found that both flipped and simulated conditions were more efficient (i.e., required less time for students to complete the material) than the traditional condition. They also found that students in all three conditions learned the material, but that students in the stimulated condition had lower learning gains than either of the other conditions, which did not differ from one another. The same pattern of results was also true for several self-reported measures, including course evaluations and survey questions such as "How valuable was this course?" and "How likely are you to take a similar course in the future?." In each case, students in simulated courses reported lower scores than those in either flipped or traditional classrooms, which did not differ from another. The researchers suggested that a specific advantage of the flipped curriculum design was the flexibility it allowed students in setting their own pace. They concluded that "the evidence suggests that the flipped approach is at least as effective as the regular approach for delivering this class and somewhat more scalable" (p. 577).

Thus, there is limited but growing support for the possible use of a flipped curriculum to enhance student learning, customize course structure and delivery, and make better use of instructional resources. However, there remains significant research to be done, particularly regarding the applicability of flipped curriculum designs to new domains, such as legal and pre-law education.

Legal Education

The explicit aim of a Law School Admissions Test (LSAT) prep course is, of course, to prepare students for the exam itself. A secondary aim of my own instruction has been to help students acquire some of the basic reasoning skills they will need for law school and other legal education. Legal education has recently come under fire for various reasons. Academic and non-academic observers alike have argued that the value of a law degree has decreased as the cost has increased, whereas others are concerned that access to legal education remains unequal, favoring white applicants over non-white applicants and male students over female students (Stake, 2006). Simultaneously, the National Association for Legal Placement estimates that there are 60,000 fewer legal jobs available in the U.S. then there were in 2007. In response to these factors, the number of applications to law schools has declined dramatically since 2010. There were 87,900 applicants to American Bar Association (ABA) accredited law schools in 2010 and 59,426 applicants in 2013, the most recent year for which the full number of applicants is available. This is a decrease of 32%, and the numbers continue a downward trend.

Furthermore, some observers have argued that current applicants are increasingly less well-prepared for the kind of critical reasoning and problem-solving required for a law school education. Stuart and Vance (2013) wrote, "Today, more students enter the legal academy without even rudimentary problem-solving skills. Indeed, emerging empirical evidence reveals that fewer students possess the basic higher-order cognitive process that the academy has assumed are the threshold educational achievement for success in law school" (p. 4). They cited the recent creation of a required, first-year

seminar on problem-solving at Harvard Law School as evidence that even presumptively elite law school students are in need of remedial critical thinking instruction.

Pre-law students. Very little research has been done on the learning outcomes among pre-law students before entering law school. Perhaps one reason this population has attracted little scholarly attention is the prevailing belief that law school applicants are a homogenous group with little of interest to offer educational researchers. Pre-law students are often perceived to be "recent college graduates, upper middle-class white male in the top 15% of their undergraduate class with a pre-law, political science, or liberal arts major" (Randall, 1995-1996, p. 68). Since 1990, the first year in which the current LSAT form was used, the pattern of law school enrollees has changed significantly (see Table 1):

Table 1. *Proportion of JD student enrollment by gender and minority status.*

	1990-91 totals	%	2012-13 totals	%
Total JD Students	127,261	n/a	139,055	n/a
Male	73,164	57.5	73,668	53
Female	54,097	42.5	65,387	47
Minority	17,330	13.6	35,914	25.8

Note: Data from www.lsac.org/lsacresources/data/ethnic-gender-admits.

In just over 20 years, the ratio of female to male students has increased substantially, and the proportion of self-reported minority students has almost doubled. Of note is the fact that although total enrollment has declined from its peak of 147,525 in 2010-11, 2012-13 marked the highest ever number (and proportion) for minority enrollment. Although no similar statistics are available for pre-law students in general, continued controversy over the effect of LSAT scores on minority admissions suggests that law school applicants

may be significantly more diverse than law school students (Cross & Slater, 1997; Randall, 2006).

Also, according to Law School Admissions Council (LSAC) figures compiled for the 2007-2008 admissions cycle, there were at least 450 law school applicants from each of 29 major groupings of students' undergraduate majors. These groupings were collapsed from all reported majors by Nieswiadomy in order to determine the effect of major choice on LSAT performance (available at Social Science Research Network: http://ssrn.com/abstract=1430654 or http://dx.doi.org/10.2139/ssrn.1430654). The ten most common major groupings included traditional pre-law categories such as political science (19.9% of applicants), English (7.7%), and liberal arts (5.1%), but also included business management (6.1%), journalism (5.7%), and psychology (4.5%). This broad range of applicants is also reflected among students who choose to engage in commercial LSAT test preparation. Additionally, an increasing proportion of law school applicants (and, by extension, of test prep students) are applying to law school after several years in a different career field. For these reasons, research regarding test preparation among LSAT students may have broader generalizability than initially assumed.

Another possible explanation for the lack of research about pre-law students is that such students are seldom together in a learning environment. Precisely because pre-law students come from a wide range of majors, there is seldom an opportunity to study learning outcomes within this population. With the exception of schools that offer criminal justice degrees, universities and colleges do not prescribe a preferred pre-law curriculum. Pre-law activities typically occur outside of classroom settings, in pre-law

societies, at law fairs, or in LSAT prep courses. Thus, nearly all of the research available for this population takes place among law students rather than pre-law students and does not include students who fail to gain admission to any law school.

However, although educational research among law students may not reflect all segments of the pre-law student population, it may nonetheless be useful in understanding the beliefs and attitudes of pre-law students, particularly among those prepared so as to be successful in gaining admission. In a qualitative study of first-year law students' beliefs about learning, Harris (2003) wrote that "law students may be expected to be almost ideal students: self-motivated, self-directed, and focused on learning activities that allow for the development of meaningful understanding" (p. 7). She added that "the cognitive learning strategies that first-year law students select may depend on their unique patterns of learner characteristics (e.g., expectations for success, personal preferences for thinking deeply), their personal beliefs about the nature of knowledge and learning (e.g., that teachers are the sole authority, that learning should happen quickly), and their approaches to learning (e.g., using simple memorization strategies or strategies that facilitate deeper learning)" (p. 14).

The LSAT. As with most graduate and professional examinations, the LSAT is a standardized exam with normalized scores adjusted to the measured characteristics of the student population. Because the LSAT is weighted so heavily in admissions decisions, scores on the LSAT may determine whether or not students are admitted to any school, in addition to determining eligibility for scholarships and selective programs (Klein & Hamilton, 1998). The test administrations are typically held in February, June,

September, and December of each year. Each administration consists of four scored multiple-choice sections, an unscored multiple-choice section used to test items for future administrations, and an unscored writing section. Each section lasts 35 minutes, and a 15-minute break is offered between the third and fourth sections.

The scoring scale consists of about 101 questions from the four scored sections (two Logical Reasoning sections, one Reading Comprehension section, and one Analytical Reasoning section). Raw scores are converted to a normalized scale that ranges from 120-180 (mean ≈ 151, SD ≈ 7.5, SE ≈ 2.7). Scores are used by law schools in various ways during the admissions process. One common use is to regress LSAT scores and undergraduate GPA (UGPA) onto first year law school GPA (FYGPA). Technical reports conducted on behalf of LSAC show that such regression equations explain between 40 and 50% of the variance in FYGPA for admitted students (Evans, Sweeney, & Reese, October 2011). Of these two factors, LSAT has a significantly higher beta coefficient than UGPA. In an analysis of the admissions formulas used by the top 100 North American law schools, Didier (2013) determined that the LSAT had a mean coefficient weight of nearly four times the UGPA. The LSAT/UGPA coefficient weight ratios Didier derived ranged from a low of 1.67 for the University of California, Berkeley School of Law to a high of 7.61 for the University of Buffalo-SUNY School of Law.

On average, more than 120,000 students take the LSAT each year. The number of applicants peaked in the 2009-10 application year, when there were 171,514 test-takers and has declined since. There were only 129,958 test-takers in 2011-12, and in 2013-14 the number dropped again to 105,532. Most law schools require a valid LSAT score, a

full transcript of all college courses taken, a personal statement, and at least two letters of recommendation. Application materials are compiled by the Law School Admissions Council (LSAC) and sent to law schools as requested by applicants. Approximately 35% of test-takers earn admission to law school each year.

LSAT preparation classes. Although there are a large number of regional and local boutique firms offering LSAT prep services, only six companies have a significant national presence (listed in approximate order of number of students): Kaplan, Princeton Review, TestMasters, PowerScore, Blueprint, and Manhattan Test Prep. Each of these companies offers a variety of courses and services, including live classes, online classes, publications, and individual tutoring. In addition to offering largely similar services, these companies also have similar curricula, a fact due in part to the nature of the test and in part to a common "family tree" of each company's founders. Kaplan is the oldest and largest of the test prep companies. Each successive national company, beginning with the Princeton Review, was founded by one or more individuals who had been previously employed at an existing company (e.g., TestMasters was founded by former Princeton Review employees, whereas PowerScore and Blueprint were both founded by former TestMasters' employees).

For each company, the current model for hiring instructors consists of soliciting instructor applicants who meet or exceed a predetermined cutoff on an officially administered LSAT examination. Depending on the company, these cutoffs range from the 95th to 99th percentile. Applicants are then interviewed, invited to audition, and, if hired, trained on the company's specific methods and curriculum. The nature of this

process, especially the initial selection criteria, promotes the belief that instructors are content experts. This belief can be reinforced by student behavior in the classroom. For example, some students expect or, occasionally, demand, that instruction consist of an uninterrupted flow of advice, tips, and explanations from instructor to student. For example, during my own classes, more than one student has criticized me for making a joke or making a brief reference to current events during a four-hour class rather than "teaching" without pause.

Critical thinking and the LSAT. The LSAT is explicitly intended to test critical thinking and reasoning skills. A recent study found evidence that intensive LSAT preparation can affect the structure of regions within the brain associated with logical thought. Mackey, Whitaker, and Bunge (2012) compared fMRI scans of undergraduate students' brains before and after a test prep course to similar scans of their peers enrolled only in undergraduate courses. The researchers concluded:

> Both groups consisted largely of university students, and their academic experiences over the course of 3 months alone could have altered their white matter microstructure. Thus, changes that were significantly greater in the trained group (test prep students) than in a well-matched control group provide strong evidence for experience-dependent plasticity, and not simply maturational changes. (2012, p. 7)

Thus, although both groups showed signs of increased myelination (the physical process of reinforcing neuronal connections) in brain regions associated with cognitive development, the test prep students showed additional development in various regions thought to be linked with complex reasoning processes.

Although many forms of critical reasoning are present on the exam, one of the most commonly tested is that of propositional logic (conditional reasoning). This concept

consists of understanding the proper relationship between sufficient indicators and necessary conditions. Nearly 25% of all questions asked on the LSAT involve some degree of conditional reasoning. My study paid particular attention to students' learning of conditional reasoning as an insight into their critical thinking development.

Research regarding experimental measures

Measure of student engagement. One focus of interest in this study is the effect of a flipped curriculum design on student engagement. As Skinner, Kindermann, and Furrer (2009) indicated, researchers from numerous fields, including assessment, teacher development, educational technology, and even neuroscience, have become increasingly interested in classroom engagement. One reason that engagement is viewed as a fertile field for further research is that it "represents a potentially malleable proximal influence shaping children's [and other students'] academic retention, achievement, and resilience" (p. 494). Additionally, Svanum and Bigatti (2009), who observed 235 students for a five-year period, noted that "engaged students were more likely to succeed than their less engaged counterparts; succeeded more rapidly; and performed better than expectations based upon" GPA or admissions exam scores (p. 128). While these results could indicate that successful students are more engaged rather than engaged students being more successful, they nonetheless justify further investigation of engagement in terms of its openness to intervention and potential for meaningful impact.

Astin (1984) defined *engagement* as "the amount of physical and psychological energy that the student devotes to the academic experience" (p. 297). Although it may be possible for a student to be engaged without one or the other of these components, typical

student engagement results in or derives from physical, emotional, and mental involvement. Accordingly, Skinner, Kindermann, and Furrer (2009) envisioned engagement as having both an emotional and a behavioral axis. For each of the two axes, the positive affect is denoted as engagement and the negative aspect as disaffection, yielding four separate but highly interrelated measures of engagement: (a) behavioral engagement, (b) behavioral disaffection, (c) emotional engagement, and (d) emotional disaffection (Skinner, Kindermann, & Furrer, 2009, p. 500). Their self-report data included responses on items intended to measure each of these components. In addition, teachers can observe many of the behavioral components among students, such as focusing on the teacher, remaining on task during class-work, attending classes, and limiting conversation to course-related topics. These observations should generally correlate with the self-report.

For my purposes, it was important to identify a reliable and valid measure of short-term student engagement to be adapted for use in my study. Many useful measures are aimed at understanding engagement on a "macro" level. The National Survey of Student Engagement (NSSE), conducted by researchers at Indiana University, measures "whether an institution's programs and practices are having the desired effect on students' activities, experiences, and outcomes" (National Survey of Student Engagement, 2000, p. 1). However, as Handelsman et al. (2005) wrote, "The NSSE focuses on active learning and other educational experiences but does not focus on individual courses; rather, it assesses students' overall perceptions" (p. 184).In order to develop a better assessment of short-term engagement, Handelsman et al. (2005) developed a scale for use during a

single course. The Student Course Engagement Questionnaire (SCEQ) consists of 24 items that loaded onto four primary factors. These factors have been deemed Skills Engagement, Emotional Engagement, Participation/Interaction Engagement, and Performance Engagement. Handelsman et al. validated the SCEQ by comparing it to constructs previously associated with student learning and motivation. They found that "all four of the SCEQ factors were associated with at least one other measure; the different patterns among the variables supported the distinctiveness of the student engagement factors" (p. 189). Additional reliability and validity measures confirm the usefulness of the SCEQ as a tool for measuring short-term student engagement. An adapted version of the SCEQ was used in this study.

Measure of team effectiveness. Given that that the primary learning outcomes for this course were measures of individual learning, I sought a measure that would offer insight into the students' perception of how effective the groups were. I considered team member evaluation as recommended by Michaelsen and Sweet (2008), but wished to focus on the teams as a whole rather than on individual members. After reviewing relevant literature regarding team-based learning, I decided to adapt the Team Learning and Beliefs Questionnaire developed by Van den Bossche, Gijsalears, Segers, and Kirschner (2006). These researchers conceived of collaborative learning as "the creation of mutually shared cognition" and sought a model that accounted for both the cognitive and social contexts of collaborative learning. Furthering research by Roschelle and Teasley (1995) and Barron (2003), they argued that the effectiveness of team learning depends not only on opportunities for collaboration but also coordinated construction of

meaning and joint navigation of interpersonal contexts. Thus, the Team Learning and Beliefs Questionnaire captures data regarding cognitive processes, such as construction and co-construction, as well as social process, such as interdependence, task cohesion, and group potency. The researchers proposed a model in which interpersonal beliefs would directly affect team learning behavior, which in turn affected mutually shared cognition and team effectiveness.

The questionnaire was administered to first-year students in two courses of an international business degree program who worked in teams of three to five students over a seven-week period. For one course, the teams were self-selected and for the other, they were chosen by the teacher. Data from 75 teams supported researchers' model, while multiple regression analysis also showed a direct effect of beliefs regarding group potency and task cohesion on team effectiveness. It was hoped that the same model would provide insight into the collaborative learning experience of the subjects here.

Conclusion

Although substantial research supports the application of the principles of active learning through a flipped curriculum method and small-group problem-solving, the particular nature of the learners, context, and content in an LSAT classroom has seldom, if ever, been investigated previously. The current study not only was meant to add to our understanding of how pre-law students learn critical reasoning skills in a self-selected and largely self-directed setting, but also to offer suggestions to other instructors about adopting active learning principles into their classrooms.

Chapter Three: Method

Statement of Purpose

The purpose of this study was to evaluate an instructional intervention designed to enhance student learning. Specifically, this project evaluated an instructional setting utilizing a flipped curriculum, small group collaboration, and student response systems. The study was a mixed-method investigation of three courses learning a critical reasoning curriculum in an LSAT preparation classroom.

Design Experiments

Methodological note. According to Reinking and Bradley (2008), the first criterion for a design experiment is that such experiments involve an intervention in authentic instructional contexts. The current study involved a number of pedagogical changes in an authentic classroom environment. Reinking and Bradley (2008) wrote, "Education researchers conducting a formative or design experiment...attempt to bring about positive change in education environments through creative, innovative, instructional interventions grounded in theory and guided by systematic data collection and analysis" (p. 6). The staggered nature of the courses allowed me to monitor and adapt the instructional design as needed between Course 1 and Course 2 (Course 3 was added subsequent to the original to provide additional sample size). I provided a multi-faceted view of the teaching and learning processes in the classroom by gathering several types of data. Some of the data were compared to traditional courses previously taught by me.

Second, design experiments should be theoretical in nature. More specifically, Reinking and Bradley (2008) wrote that design experiments "are also aimed at

theoretically understanding the conditions that enhance or inhibit an intervention's effectiveness and at generating pedagogical understandings that can generalize beyond specific instances" (p. 19). This study examined the application of collaborative learning and problem-based learning theories among a group of adult learners in a self-guided learning environment. Among the conditions of interest here were whether the typical benefits of collaborative and problem-based learning extend to a learning context with potentially competitive learners and when the instructor has none of the usual means of enforcing accountability. Third, design experiments should be goal oriented, with the explicit purpose of improving an educational setting. Again, this criterion was met by my study, as one of my goals was to design and describe an instructional approach that would allow these students to learn better the critical thinking and reasoning concepts tested on the LSAT and required for success in law school.

Participants

As approved by the Institutional Review Board, the experimental condition consisted of students in three courses of an LSAT prep course, one beginning in July 2013 (Course 1), one in August 2013 (Course 2), and one in October 2013 (Course 3). Course 1 had 18 students (one of whom transferred to another course location after lesson 2), whereas Courses 2 and 3 each had 11 students. Thirty-seven of the initial 40 students consented to be included in the research. The students were 20 males and 20 females, ranging in age from 19 to 40 years old. Further demographic data are provided in Tables 8 and 9 in Chapter Four.

Informed consent was obtained electronically from all participants following the

first course meeting. During that meeting, students in the flipped curriculum condition were told the purpose and format of the study (data from control students in previous courses were strictly aggregated and did not require informed consent). Following the first meeting, I sent students a link to the first survey, which contained the informed consent form. Only I had access to the survey results, and all students had anonymous identifiers.

Materials and Measures

Materials. Materials for this study included a classroom, a laptop, student response devices (typically, the students' own laptops or smartphones), twelve LSAT lectures (each with homework problems, drills, and supplemental materials), four official LSAT exams, and ten supplemental video lectures.

Measures. Students taking the courses with a flipped curriculum were asked to take three surveys, one after practice test 1, one after lesson 6, and one after lesson 12. The control group was not asked to complete any of these surveys, as the control data were retroactive and aggregated.

As indicated previously, the first survey contained the informed consent for this study. It also gathered basic demographic data, information about students' study goals, and selected results from their first practice test (specifically, their scores on six conditional reasoning questions measuring Logical Reasoning).

The second survey asked questions about the students' experience with group learning. Specifically, it contained 35 items adapted from Van den Bossche, Gijsalears, Segers and Kirschner's (2006) Team Beliefs and Behaviors Questionnaire. These items

were scored on a five-point Likert scale and addressed student experiences in several domains, including interdependence, social cohesion, group potency, and co-construction. Below are the items for interdependence, the responses to which are averaged together to yield the subscale score (for a list of all items, see Appendix B):

Table 2. *Team Learning Beliefs and Behaviors Questionnaire subscale measuring team interdependence.*

1.	My team members depend on me for information and advice.
2.	I depend on my team members for information and advice.
3.	The team members agree on what we want to accomplish.
4.	When my team members succeed in their jobs, it works out positively for me.

Note: Adapted from Van den Bossche, et al. (2006).

This survey also provided students with an opportunity to describe their experience with group learning during the first half of the class. Responses to this survey were used to adjust the implementation of group work during the second half of Course 1, and throughout Courses 2 and 3.

The final survey included a modified version of the Semester Course Engagement Questionnaire (SCEQ) (Handelsman, Briggs, Sullivan, & Towler, 2005), as well as six conditional reasoning questions from the students" initial practice test. As described in Chapter Two, the SCEQ is a six-point Likert-type survey that elicits responses regarding both in-class and out-of-class interaction with course material, concepts, and other students. Below are a few sample items from the SCEQ (for a full list of items in the SCEQ, see appendix C):

"I make sure to study for this class on a regular basis."

"I put forth effort in this class."

"I do all of the homework problems for this class."

"I look over class notes between classes to make I understand the material."

"I take good notes in class."

The students' responses to the conditional reasoning questions were compared to their responses to the same questions on Test 1, which was taken prior to receiving any course instruction.

Another measure was the comprehension quizzes administered at the beginning of most classes. These quizzes served several purposes: First, the quizzes tracked student attendance and homework completion. Second, the quizzes determined overall student comprehension, which, in turn, determined what topics, if any, I reviewed prior to group work in each class. Finally, the quizzes were used after each practice test to gather student test scores.

Process

To understand how the classroom design was changed during this study, it is important to understand the traditional course structure. As a brief overview of the traditional pedagogical approach was provided in Chapter One, I focus here on a more detailed description of a typical lesson.

Traditional instruction. Lesson 1 covers two logical reasoning question types (Must Be True questions and Main Point questions) and one category of logic games (Sequencing games). Students have already taken a practice test prior to lesson 1, but are not assigned homework or review exercises. The lesson begins with a discussion of the section directions and scoring procedures for logical reasoning sections on the LSAT. Special note is made of the manner in which the LSAT treats prior knowledge and the

testmakers' approach to writing incorrect answers. The instructor lectures about the main parts of a logical reasoning question (stimulus, question stem, and answer choices) and recommends that students not attempt to game the test by reading the question stem or answer choices first.

Next, the instructor demonstrates PowerScore's recommended approach by modeling a sample logical reasoning problem (instructors occasionally seek student participation at this point and will ask a student to read the problem). Using this question as a model, instructors discuss conclusions and premises, along with typical indicator phrases for each. In the past, I have typically added that students should focus on the relationship between the premises and conclusion, and that they are largely expected to accept the premises as stated. Four categories of critical language are discussed, before the instructor briefly introduces each of 13 question types.

A key concept of PowerScore's approach is the method of addressing answer choices. The instructor teaches students to read all the answers before picking any, separating plausible answers (Contenders) from implausible ones (Losers). Students are also taught that confusing answers should automatically be considered plausible and that none of the answers should be considered in-depth until all of the answers have been at least addressed. After briefly reviewing the key concepts for logical reasoning, the instructor introduces Must Be True questions.

Students are taught to rely only upon information contained within the stimulus and to accept all of this information. The key question for this problem type is, "If I believe everything in the stimulus, what else must I believe as a result?" Students are

taught to reject true answers that are not drawn from the stimulus, answers that are too extreme for the information from the stimulus, answers based on information external to the stimulus, and answers that change a key word from the stimulus. The instructor typically guides students through several practice problems before challenging the students to attempt a question on their own. Guided questions consist of a student reading the stimulus aloud for the class, and the instructor asking questions to the class to draw their attention to critical elements of the stimulus. The student then reads each answer choice aloud, as the instructor asks students to rate each answer as a Loser or Contender. Students who deem an answer to be a Loser are invited to explain their rationale, whereas answers considered to be Contenders are saved for discussion until after all the choices are read. Practice questions performed by students on their own during class are typically timed. Although student responses are not formally collected, instructors often wander the room to see what answers are commonly chosen and frame their post-question explanation around this informal process.

Lesson 1 contains 10 Must Be True questions, of which the instructor is typically expected to cover at least seven during the lesson. It is at the instructor's discretion to determine how many questions will be guided and how many will be individual practice. The logical reasoning section of lesson 1 concludes with a brief overview of Main Point questions, and two sample problems. Students are given a 10-minute break and continue with the logic games section. The overall balance between lecture and student practice is similar to that in the logical reasoning section.

Flipped instruction. A goal of flipped instruction is to increase time on task

without reducing the amount of content covered in a class. Accordingly, one goal for this intervention was to provide students with the same lecture content as traditional instruction by asking them to watch a video before coming to class. For lesson 1, I recorded two videos, one for the logical reasoning section and one for the logic games section. In the video, I instructed students to read the lesson text before proceeding and then to follow along with the book while watching the remainder of the video.

The logical reasoning video for lesson 1 was 18 minutes long and generally followed the same content as the traditional lecture. However, because I presumed that students had already read the lesson text, I emphasized slightly different aspects of the material. The similarities and differences to the traditional lecture are described below. The video began by discussing the section directions for logical reasoning, and then addressed the rationale for always reading the stimulus first. I emphasized how reading the question stem could lead to confirmation bias. I then discussed the nature of argumentation, and defined premises and conclusions with respect to the LSAT. I compared premises to evidence that had been admitted to trial during discovery and must be accepted by students. I explained that students are seldom allowed to attack premises and are to focus instead on questioning the connection between the premises and the testmakers' conclusion.

I then contrasted standard reading with LSAT reading. In standard reading, we pay most attention to the "meat" of an argument (specific facts or concepts), but the LSAT is most concerned with the "gristle" of an argument (connective words or phrases, adjectives, verbs, and modifiers). On the LSAT, the specific facts or elements of an

argument are secondary to the roles played by each element. Students are warned not to simply spit out the argumentative gristle.

As in the traditional lecture, I discussed the features of various question types and the roles played by different kinds of answer choices or stimuli. After covering the categories of question types, I spoke briefly about prephrasing answers with specific but flexible guesses about what the correct answer must do. In discussing answer choice evaluation on the video, I spent several minutes comparing the Losers and Contenders approach to a high jump competition. Students were told that prior to reading any answers, their standard for picking a contender should be relatively lenient, similar to the qualifying height in the first round of high jump competition. The goal of having a low bar is to reduce the possibility of hastily eliminating correct answers. I explained that as students encounter better and better answers among the five choices, they should continue to raise the bar and set a higher standard for future answers. Each new contender should be at least as good as the best answer already encountered. I also explained that confusing answers should be set aside and kept as contenders.

Students were reminded during each class to watch the video prior to the next lesson and not to do any of the lesson questions before coming to class. At the beginning of lesson 1 in the flipped classroom condition, I administered a comprehension quiz via Socrative, a web-based classroom response system that allowed students to respond to questions using personal electronic devices, such as tablets, smartphones, or laptops. Depending on the lesson, these quizzes had between seven and ten items, and contained a mix of multiple choice and short answer questions. Lesson 1's quiz consisted of nine

questions (three short answer questions and six multiple choice questions). Because a significant number of students arrived after class began for each lesson, this quiz was open for the first ten minutes of class and was student-paced (i.e., each student individually controlled which questions were displayed at which pace on the classroom response system). Accordingly, I could only analyze answer choice patterns after the quiz was closed and I download the results. The questions were intended to be easy for prepared students to answer, so if more than a third of students missed a particular question, I offered a brief lecture on that concept.

After the quiz and any mini-lectures, the balance of the class was spent in collaborative problem solving. Five of the thirteen lessons in Course 1 took place in a room with tables that allowed students to face each other. For these lessons, the groups were pairs of students sitting at the same table rather than larger groups of three or more students. However, as I noticed any pair of students struggling to understand a concept, I invited them to work with the students seated directly in front of or behind them. Thus, the actual group size ranged from two to four students, and there was some variation as to which students sat next to each other. The remainder of lessons in Course 1 and all of the lessons in Courses 2 and 3 took place in a room with stadium-style seating, forcing all students to face forward. During these lessons, groups were generally three to four students each.

The amount and nature of group work in each lesson of the flipped course varied by section. The general procedure for questions from the logical reasoning section is described here. During the group work portion of each lesson, students were asked to

solve one or two logical reasoning problems individually, answer the problems through a Socrative quiz, and then discuss their answers and overall approaches with their partners. As students discussed the problems, I circulated the room, looking for concerns, questions, or other opportunities to assist them. I also prompted each group to discuss more than just the correct answer. For example, I asked them to find the best incorrect answer and articulate exactly what convinced them not to pick it. When students disagreed about the correct answer or the best incorrect answer, I asked them to explain their reasoning to each other, but did not insist that they reach consensus. After a few minutes of discussion, I reconvened the class to explain any pertinent features of each problem.

For the logic games section, each traditional lesson includes a lecture and three or four practice games. The first one or two games are models demonstrated completely by the instructor and the remaining games are used for individual student practice. The flipped curriculum design required the students to view a lecture and read the lesson text before class and instruction began with a model game. After demonstrating an appropriate set-up and working through all of the questions for the model game, I asked students to do a setup collectively for the next game. As the groups discussed their setups, I asked questions about their approaches and offered suggestions to help clarify their diagrams. Once all students had an acceptable set-up, they answered the questions individually. Finally, students were given the opportunity to do a game by themselves. Although the number of games completed varied from lesson to lesson, there was always at least one model game, one collective game, and one individual game. When time

allowed, I added a second collective game before the individual game was presented.

Reading comprehension was taught largely as in the traditional course, with in-class lectures and no group work. This decision was both practical and methodological; reading comprehension does not lend itself well to group work and keeping this instruction constant was intended to provide a means of distinguishing student or instructor effects from intervention effects. I compared initial reading comprehension scores and score increases for the flipped curriculum course to those of previous traditional courses.

Data Analysis

Quantitative data analysis. To measure the effect of the intervention on student learning, I compared change scores from test 1 to test 4 and also from each student's minimum score to maximum score between traditional and flipped courses. For this measure, students from three recent traditional classes were collapsed into the traditional group, whereas the flipped group consisted of students in Courses 1, 2 and 3. Test 1 scores for all classes were compared to determine if students in each course had similar levels of pre-existing knowledge.

Other measures only involved students within the flipped courses. For example, after lesson 6, the flipped course students were given a survey regarding their beliefs about teamwork and experiences with group instruction during the course. These results were used both to determine if differing levels of belief in teamwork can moderate the effect of flipped course instructions on learning and engagement, and, if necessary, to adjust the implementation of group instruction in these courses. Because lesson 6 in

Course 1 occurred before Course 2 began, the results of the survey were used to inform the implementation of group work not only in the second half of Course 1 but throughout Course 2.

Research question 1. Compared to traditional LSAT prep instruction, what effect does a flipped curriculum have on student learning, student engagement, and student course persistence?

Data for question 1 were drawn from students' test scores and selected survey questions. Primary statistical analysis consisted of t-tests between the experimental group (for these analyses, students in Courses 2 and 3 were collapsed and compared with students from Course 1), and students who had taken courses with me prior to the study, specifically 26 students who had enrolled in LSAT courses the year before the study and for whom I could construct measures of student learning and student course persistence.

Qualitative data analysis. Research question 2: What is the experience, for the instructor and students, of collaboratively learning in an LSAT prep course?

Research question 3: What does group discourse reveal about the learning environment and process within a course using a flipped curriculum?

In order to address these questions, I relied upon several forms of qualitative data. I audiotaped my own comments during each lesson and recorded each group's discussion during lessons 2, 5, 6, and 11. I used a constant comparison, emergent theory method to code for concepts, themes, and categories related to both the students' collaboration with each other and their understanding of critical reasoning concepts. I transcribed each group's audio recordings and then analyzed and coded the transcriptions to find emergent

concepts in the groups' discussion. I then openly coded these concepts to derive themes based on the participants' use of terms and concepts from the course. I was particularly interested in the students' use or misuse of specific critical reasoning terms and how they explained their thought processes to themselves and each other. Other qualitative data sources included observation field notes recorded during and after each class, short answer responses to survey questions, and interviews with a select sample of students from each course. The results of these analyses are discussed in Chapter Four.

Chapter Four: Results

In this chapter, I include both quantitative and qualitative results for this study. I had initially planned to compare the test scores and engagement questionnaires for my students with those of students in sections taught by other instructors. To that end, I contacted the headquarters of the company that administered the course I was studying, PowerScore, and asked for the names of other experienced instructors who would be teaching courses during the time I was collecting data in my class. I was given the names of seven other instructors, all of whom taught near urban areas with tier I schools (New York; Washington, DC; Chicago; Berkeley; Davis, CA; Ann Arbor; Dallas). I asked each instructor to invite his or her students to complete pre- and post-course surveys for my study. These surveys primarily consisted of demographic information, a team-based beliefs questionnaire, the SCEQ, and general questions about the students' scores, goals, and preparation.

I sent four emails to the instructors, concurrent with the beginning and end of the July and August courses I was studying/teaching myself (Course 1 and Course 2). Altogether, I received some minimal response from four of the seven instructors, indicating assent to share my request with their students. However, despite these efforts and my eventual decision to offer small gift cards to anyone who completed both surveys, only five students, altogether, completed the pre-test survey. As a result, the engagement scores reported here have no control group and can be used only for within-group analyses.

For test score comparison, I had to devise a different comparison group. Instead

of comparing these participants' test scores with those of students taught in concurrent courses by other instructors, I used archival data from my own previous courses. I selected three courses from 2012 (July, August, and October) that I had taught with a traditional curriculum (the year before data were collected for this study). The three control courses and the three flipped courses all used the same practice tests, lesson texts, and practice problems. All students received the same number of course hours, and all six courses were taught by the same instructor, me, in the same location. However, before comparing the three flipped courses to the control courses, I first determined if there were any differences among the flipped courses.

Quantitative Results for Flipped Courses

Much like my attempts to gather control data from other instructors, I faced significant difficulties gathering data from my own students. The course offers four exams that are strongly recommended to students as practice and then self-scored in an online student center. Despite my emphasis on the importance of these exams for the students' preparation and my repeated requests for students to share the results with me through the student center, many of the students did not complete these tests. Of the 38 students who took test 1 in any of the flipped courses, only 22 took test 4, and only 15 students took all four tests.

In addition to the four practice exams, my other primary source of quantitative data was the initial, mid-course, and post-course surveys (described in detail in Chapter Three). Links to each survey were posted during class and emailed to students in each of the three courses. The initial survey was emailed to students following the introductory

lesson, and the mid-course and post-course surveys were distributed following lessons 6 and 11, respectively. Despite repeated requests and my offer of $10 gift cards to students who completed all three surveys, only 15 of the 38 students completed all of the surveys, and five students did not complete any of the surveys. These students were similarly reluctant to provide data following the course. During my data analysis, I emailed ten students who had not yet responded to ask for basic demographic information, and received only one reply.

Overall test scores and attendance. Because the participation rate was far lower than I had anticipated, any quantitative analyses grouped by course had very little power, and there were few significant differences between the courses. After careful consideration, I determined that the primary design iterations during my study had occurred between Courses 1 and 2, and that the intervention was essentially unchanged for Course 3. Therefore, I collapsed Courses 2 and 3 in my statistical analyses (hereafter referred to collectively as Phase II), and tested for differences with course 1 (hereafter, Phase I) on test score increases. The findings are reported in Table 3.

Table 3. *Mean Score Increase from Test 1 to Highest Test for Flipped Courses.*

	Phase			
Score Increase	I (n = 14)	II (n = 20)	*df*	*t*
Overall Score	9.1 (3.8)	7.1 (4.5)	32	1.35
Logical Reasoning (%)	16.9 (7.5)	12.7 (10.3)	32	1.31
Reading Comprehension (%)	16.2 (7.4)	8.1 (17)	32	1.69
Logic Games (%)	21.6 (15.7)	21.3 (15.2)	32	.05

Note. Standard deviations appear in parentheses below means.

Because so many students failed to take test four, the gain scores were calculated by comparing the initial scores to the highest available score from any other test (students who took only one test were excluded). Overall scores are reported on the standard LSAT scale of 120-180, and logical reasoning (LR), reading comprehension (RC), and logic games (LG) section scores are reported in percentages of questions answered correctly. Thus, although students in Phase I gained nine points (compared to seven points for students in Phase II) and improved in reading comprehension by 16% (compared to 8% for students in Phase II), none of these differences was significant.

I also compared average attendance for the flipped courses. In traditional courses, the average attendance has been 79% (as defined by the actual number of students attending each class divided by the number of students enrolled) and in the flipped courses, the attendance rate was 75%. Interestingly, the attendance for Course 1, which had 18 students enrolled, was only 67%, making it the worst attended course for which I

have taken attendance. Courses 2 and 3 had higher than average attendance (80% and 83%, respectively). T-test results for the effect of phase on overall average attendance are in Table 4 (representing number of classes attended out of 13 total and the sample includes, as the denominator, all students who attended any session during the course):

Table 4. *T-test for Mean Attendance between Phase I and Phase II.*

	Phase			
	I (n = 18)	II (n = 23)	*df*	*t*
Attendance	8.3 (2.9)	10.5 (2.3)	39	-2.64*

Note. *p < .05. Standard deviations appear in parentheses below means.

There was a significant difference in attendance between phases, with students in Phase I attending fewer classes than students in Phase II.

Furthermore, some students in all courses, particularly Course 1, often left class early. On several occasions, students would ask if we were "only doing reading comprehension" after our break before deciding whether to leave. Once again, course 1 was something of an outlier in this respect, with one student deliberately skipping several reading comprehension sessions, and two or three other students doing so somewhat less regularly. During lesson 10, four of the eight students left class by the midway point of the lesson. No other class in this study (or in my prior ten years of experience) had so many students demonstrate this pattern of selective attendance.

Team Beliefs and Behaviors Questionnaire. At the midpoint for each of the three courses, I administered a Team Beliefs and Behaviors questionnaire to the students. The response rate for this survey was 61% for Phase I and 53% for Phase II. Eleven students from Phase I and 13 students from Phase II filled out the survey. One student indicated "I agree completely" for all 35 questions, including three reverse-coded questions, and was therefore removed from the analysis. A second student indicated "I disagree somewhat" for the first two questions, and then "I agree completely" for the remaining 33 questions, including all three of the reverse coded-questions and this student was also removed from the analysis. A third student answered "I agree completely" to all of the standard coded questions, but answered "I disagree completely" to all of the reverse coded questions, and was therefore retained in the analysis. Average ratings for each item were collapsed by factor, and a table of descriptive statistics and t-test results for mean differences between Phase I and Phase II is provided in Table 5.

Table 5. *Mean Scores by Factor for Team-Based Beliefs and Behaviors Questionnaire*.

	Phase			
Teamwork Factor	I (n = 11)	II (n = 13)	*df*	*t*
Interdependence	4.07 (0.67)	3.94 (0.37)	22	0.59
Social Cohesion	4.18 (0.65)	4.15 (0.71)	22	0.1
Task Cohesion	3.91 (0.98)	3.85 (1.08)	22	0.15
Construction	4.03 (0.71)	4.13 (15.2)	22	-0.53
Co-construction	4.06 (0.83)	3.9 (0.9)	22	0.46
Conflict	4.0 (0.93)	4.03 (0.83)	22	-0.07
Shared Cognition	3.82 (0.96)	4.69 (0.52)	22	-2.84**
Team Effectiveness	3.97 (0.97)	4.0 (0.82)	22	-0.08

Note: **p < .01. Standard deviations are shown in parentheses below means.

In Phase I, students chose their groups, and the groups remained relatively stable across lessons, whereas groups in Phase II were randomly assigned and varied for each lesson. The results suggest that students in both Phases perceived groups to be highly functional, with the lowest factor score at 3.82 (out of a possible 5).

Also, the different group composition approaches (student selected in Phase I and randomly assigned in Phase II) had no observed effect on student perceptions of group quality for all but one factor. There was a significant effect of phase on shared cognition, as measured by questions regarding the teams' common understanding of the task and the

process for group work. During Phase I, particularly at the beginning of the course, students from several teams expressed confusion and concerns about the purpose of group work. As a result, I presented more explicit instructions and a clearer rationale for using group work during Phase II compared with Phase I. Thus, the fact that students in Phase II felt that their groups better understood what was expected of them is likely due to a difference in my instruction rather than a difference in the groups themselves.

Student Engagement. Phase I and Phase II students were also given a modified version of the Student Course Engagement Questionnaire (Handelsman, Briggs, Sullivan, & Towler, 2005) at the end of each course. The questionnaire consisted of 14 Likert-type statements using a six-point rating scale, comprising three engagement factors: 1) skills engagement (i.e., study strategies such as "taking good notes during class" and "reviewing notes between classes"); 2) emotional engagement (i.e., student involvement with course material such as "applying course material to my life" and "strongly desire to learn course material); and 3) participation engagement (i.e., interaction with students and other instructors such as "helping other students learn the material" and "asking questions during class"). Only eight students from Phase I and eleven students from Phase II completed the questionnaire. Table 6 contains the descriptive statistics and t-test results for mean scores on each factor for Phase I and Phase II.

Table 6. *T-test of Mean Responses to SCEQ Factors for Phase I and Phase II Students.*

	Phase			
Engagement Factor	I (n = 8)	II (n = 11)	*df*	*t*
Skills	4.83 (1.09)	4.67 (0.85)	17	0.34
Emotional	4.68 (1.35)	4.91 (0.78)	17	-0.48
Participation	4.92 (0.92)	4.76 (0.75)	17	0.42

Note: Standard deviations are shown in parentheses below means.

Neither the overall factors nor any of the individual items varied significantly between Phases. For all three factors, the mean score approached 5 ("Agree moderately"). Items with a mean score above 5 for both Phases included "I put effort into class," "I listen carefully during class," "I think about the course between classes," "I strongly desire to learn the course material," and "I enjoy attending class." The lowest overall item scores, averaged across both Phases, were for "I make the course relevant to my life" (4.15), "I apply the course to my everyday life" (4.27), and, surprisingly, "I help others learn the course material" (4.27). Eleven of nineteen students rated this statement as "Agree slightly" or lower, indicating that a majority of the class felt that they contributed little or nothing to their classmates' learning.

Four additional questions regarding overall course satisfaction were administered to these nineteen students, subsequent to the SCEQ. The statements included "I would recommend this course to others" and "I feel this course was a good value", and were all five-point Likert-type items (1 = "Disagree strongly", 5 = "Agree strongly"). The mean scores for these items ranged between 4.52 ("This course met my expectations") and 4.78

("I would recommend this instructor to others"). Only one student rated any statement as less than "Agree", choosing "Neutral" for both "This course met my expectations" and "I feel this course was a good value." These results suggest that students had a generally positive affect regarding the course and instructor, and that these feelings did not differ by phase.

Quantitative comparison between flipped and traditional courses. After determining that the two Phases of the flipped course implementation did not significantly differ from each other on any learning outcomes, I then compared the whole group of students to the students from the previous year who took the course in its traditional format. Mean comparisons were calculated with all students who took at least two tests, comparing overall score change and percent change in the logical reasoning, reading comprehension, and logic games sections. Because many students did not take both the first and fourth tests, I compared students' first scores to their highest score on any subsequent test. This allowed an increase in the number of participants for each condition from 16 traditional students and 22 flipped students to 25 and 34 students, respectively, and increased the power for detecting a moderate effect (α=.05) from .31 to .46. For students in the experimental condition who only took the first two tests or scored higher on the second test than any other, there is a possible concern about whether they had been sufficiently exposed to the flipped curriculum and group work for a measurable effect to occur. However, because the second test took place after lesson 5 or lesson 6, these students had experienced between 20 and 24 hours of flipped instruction, and it is reasonable to believe that the effect of the redesigned curriculum, if any, would be

detectable by that point in the course.

On an independent means comparison of initial scores, there were no significant differences between students from flipped and traditional courses on test 1 for overall score or section sub-scores. In fact, the overall mean scores were 148.7 for students from flipped courses and 148.4 for students from traditional courses (students from flipped sections tended to have slightly higher reading comprehension scores, and marginally lower logical reasoning and logic games scores). The highest overall mean scores and highest mean sub-scores were also not significantly different (156.7 flipped courses and 154.6 for traditional courses). Table 7 displays the results of an independent means t-test comparing increases in overall score and section sub-scores between flipped and traditional students.

Table 7. *T-test of Mean Overall Score Increase and Section Sub-score Increase for Students from Flipped and Traditional Courses.*

	Courses					
	Flipped (n = 34)	Traditional (n = 26)	*t*	*df*	*p*	Mean Difference
Overall score increase	7.9	6.2	1.49	58	.14	1.91
Logical Reasoning increase (%)	14.6	10.1	1.85	58	.07	4.5
Logic Games increase (%)	21.6	11.7	2.30	58	.03	9.7
Reading Comprehension increase (%)	11.7	12.3	0.17	58	.86	-0.6

Note: *p < .05. Standard deviations are shown in parentheses below means.

These results indicated a non-significant trend of greater overall score increase for

students from flipped courses compared to students from traditional courses. A marginally significant difference was observed for logical reasoning score increases and there was no difference in reading comprehension score increases. Students from flipped courses showed greater improvement on logic games than students from traditional courses. The 9.7% difference in mean logic game subsection improvement equates to 2 to 3 more logic games questions answered correctly by students in flipped courses and to a 1 to 2 point difference in overall scores, which is consistent with the overall score trend and with the direction of score trends in the reading comprehension and logical reasoning sections between the flipped and traditional conditions.

The combination of missing data, small sample size, and the necessity of devising an ad hoc control group all limit the power and value of these quantitative results. For a design experiment, none of the results had sufficient statistical or practical significance to justify specific pedagogical changes or to elucidate the experience of students within these classrooms. Therefore, the remaining analysis will lean heavily on classroom recordings, written survey responses, interviews, and field observations to help explain more fully how the students and I perceived the flipped curriculum and use of group work in these courses.

Qualitative Results

This section details the qualitative results for this study. Demographic details regarding the students in each phase are provided (to provide confidentiality, all names have been replaced with pseudonyms), along with a summary of student responses to survey questions regarding the flipped curriculum and group work. Next, analysis of

selected group recordings is presented, along with major themes from both phases. Finally, the discussion of a single group is analyzed in-depth, and compared to the process of two other groups during the same lesson.

Phase I. Phase I consisted of 18 students, nine men and nine women. Five of these students failed to provide demographic information and two did not provide initial LSAT scores for practice test 1. Four of the students had received legal training as undergraduates, and the remaining came from a variety of undergraduate majors, including philosophy, economics, and geography. One student did not attend any classes after lesson 1, and another transferred to a different location following lesson 2. Neither of these students was present in the group conversations that were analyzed during this study. The other students, including those for whom no demographic information was available, are all present in at least one of the group conversations. Complete demographic data are presented in Table 8.

Table 8. *Demographic information for Phase I students.*

Name	Goal Score	Initial Score	Age	Gender	Undergraduate Major
Alex	N/A	N/A	N/A	M	N/A
Alberto	176	149	23	M	Philosophy
Ana	165	141	22	F	Criminal Justice
Andrea	170	160	29	F	Accounting
Barry	N/A	145	N/A	M	N/A
Bertha	N/A	141	N/A	F	N/A
Katrina	170	157	21	F	International Business & Spanish
Bill	168	146	32	M	Law L.L.B
Bonnie	160	134	25	F	Legal Studies
Bret	170	127	23	M	Geography - Water Resources, Economics
Chantal	165	146	21	F	Political Communication
Chris	167	153	22	M	International Relations & Global Studies
Cindy	170	149	31	F	Psychology
Claudette	163	140	21	F	Criminal Justice
Colin	180	156	21	M	Rhetoric & Writing
Kyle	170	153	29	M	Economics
Gonzalo	N/A	138	N/A	M	N/A
Josephine	N/A	N/A	N/A	F	N/A

Note: N/A indicates that the student did not provide this information.

Goal scores indicate the students' desired result on the officially administered LSAT, and these scores were consistent with those chosen by students in earlier courses. The lowest goal score was 160, which is the 80th percentile of all students tested. Although the range of goal scores is quite narrow, there is a moderately, marginally significant positive correlation with initial scores, $r(11) = .5024$, $p=.08$, suggesting that students were somewhat conscious of their initial aptitude when setting goal scores. However, the average desired increase was 21 points, which is far more than the average increase of 9 points in my previous classes, information that I had shared with them prior to asking

them to report their goal scores.

Flipped curriculum. During the introductory meeting for Phase I, I explained the intended curriculum and instructional design. Two students mentioned that they did not have their lesson materials yet, so I decided to delay flipping the course until lesson 2. Students were asked to score practice test 1 and be ready for a standard lecture during lesson 1. I also sent a link for all of the lecture videos to the students and to the first questionnaire. Thirteen of the 18 students completed the questionnaire, and 16 students scored practice test 1. Lessons 2-11 were flipped (lesson 12 was primarily a review session with no video and very little written lesson text), and each of those classes began with a comprehension quiz.

During each comprehension quiz, I polled students on their preparation for that day's lesson. Initially, I only asked if they had watched the videos, but by lesson 9, I added a question about reading the lesson text. Generally, students in Phase I came to class prepared, especially during the first half of the course. For lessons 1 through 4, I only recorded three instances of students who came to class on time who had not watched any of the videos for that lesson (students who were more than ten minutes late did not answer the quiz). However, five of 12 students did not watch the videos for lesson 6, and four of seven students did not watch the lesson 11 video. Nonetheless, very few students came to class having neither watched the video nor read the lesson text.

In the mid-course questionnaire, several Phase I students stated that they felt the videos duplicated course content, which may explain why many of them chose to watch fewer videos as the course progressed. Bill said, "[The videos] were useful in the

beginning of the course," and Andrea said, "I would have like[d] less overlap between class and videos to dedicate class time to a more thorough understanding of that session's topic and allow for more in-class practice questions." Others were more blunt: Katrina said, "I felt that much of the material from the lecture videos was repeated in class. This made me hesitant to watch them." And Colin said, "I did not find them very useful. Honestly, I think you would just be saving yourself a lot of time by not even putting them up." As a result of this feedback, I made a conscious effort to avoid duplicating video content in class during Phase II of this study.

Group work. Group work varied between paired collaboration and small groups of three to four students each. The groups were chosen by the students and seemed to be based on seating arrangements. Approximately one-third of the lessons took place in a room where the seating allowed students to face each other, while the remainder occurred in a stadium-style classroom with all seats facing the instructor.

The composition of groups varied based on attendance and to some extent, the students' order of arrival in the classroom. Several dyads remained largely intact throughout the course, including Claudette and Chantal, and Alex and Gonzalo, each of whom knew their dyad partner prior to the course. Katrina and Barry also worked together frequently and, by the middle of the course, were flirting with each other and generally unengaged with their classmates. Both the students and I noted that certain social relationships within student-selected groups sometimes interfered with group effectiveness. Bonnie noted that it was "harder to work well with [a] group when there are already cliques in the group" and Katrina wrote, "Since we [Barry and I] were good

friends, we often did not stay focused."

Several students suggested that the groups may have been more effective if structured differently. In an interview, Katrina mentioned that it might be helpful to divide students into groups by initial score level, so that the pace would be more uniform within each group, but quickly noted that it might be problematic when students realized how the groups were formed. Colin expressed a similar thought, along with a similar concern: "Maybe separating the class into adept vs non adept students? But that would be a little rude." Bertha also said, "I would divide people into groups by level, perhaps based on results on the initial test."

Although I had considered creating stratified groups based upon pre-test scores, I rejected this approach because I felt that differing ability levels could benefit both high- and low-level students. Unfortunately, students of higher skill levels often gravitated toward each other in the groups, at the expense of explaining concepts to students who were lower performing. Chris noted that other "group members are more competent with respect to logic games and move much faster than me." Indeed, in listening to the recordings of Chris's groups, I observed several occasions where other group members seemed unaware that he had not grasped a concept and only paused to explain when I prompted them. Chris added that he felt that "teamwork should be more focused on each singular member of the group understanding the material."

During Phase I, I relied extensively on the Socrative classroom response system, a web-based electronic feedback program, to decide when students should work in groups. My approach for teaching the logical reasoning section was loosely modeled on Mazur's

(1997) peer instruction techniques in varying the amount of student-student interaction based on the proportion that missed each question. If all or nearly all of the students had answered a question correctly, I asked students to explain briefly their answers to each other before leading a class-wide discussion. If the responses indicated wider confusion (generally at least 33% incorrect responses), I spent additional time with each group individually to understand better their specific misconceptions before addressing the class as a whole.

Perhaps in part because the students were using their own laptops and cell phones as the classroom response devices, there was a pronounced tendency in this class for students to engage in texting, web-surfing, or using apps during discussion and instruction. Because the students were customers of private test preparation company, I did not attempt to ban diversionary use of electronic devices. However, the trend in this course prompted me to reduce the use of the Socrative system during Phase II.

Phase II. The students in Phase II were enrolled in two courses, one aimed at preparing students for the October 2013 LSAT and the other, for the December 2013 LSAT. Below are the 22 students from these courses for whom I have test results or survey data. Eleven students were enrolled in each course, and all students were asked to complete a demographic questionnaire following the introductory session. The demographic data are presented Table 9.

Table 9. *Demographic information for Phase II students.*

Name	Goal Score	Initial Score	Age	Gender	Undergraduate Major
Danielle	165	136	21	F	Public Relations
Debby	170	153	22	F	Government
Dolly	170	159	24	F	Art & Performance
Don	165	151	21	M	Economics/Government
Nate	170	154	24	M	English
Earl	N/A	144	N/A	M	N/A
Tobias	170	145	26	M	Interdisciplinary Education
Ernesto	N/A	138	N/A	M	N/A
Fay	180	159	22	F	Plan II Honors/Rhetoric and Writing
Fernand	N/A	144	N/A	M	N/A
Fiona	175	153	19	F	English
Franklin	174	151	24	M	Anthropology/Spanish
Fred	163	153	22	M	Finance
Gabrielle	175	161	20	F	English
Gaston	168	157	32	M	Business and Economics
Gordon	N/A	N/A	N/A	M	N/A
Grace	165	149	21	F	Marketing
Hanna	173	156	24	F	Music
Harvey	160	144	33	M	Business Administration
Helene	N/A	146	N/A	F	N/A
Ida	N/A	142	N/A	F	N/A
Nicole	172	159	40	F	Business

Note: N/A indicates that the student did not provide this information.

Similar to students in Phase I, these students set very high goals, ranging from 160 to 180. Once again, the range of goal scores is narrow, and there is a moderate, positive correlation with initial scores, $r(16) = .60$, $p < .05$. These students were also highly optimistic about possible improvement, seeking, on average, an increase of 17 points.

Flipped curriculum. Students in Phase II were generally prepared for class. Although I began polling students about their compliance with reading course material

and viewing lecture videos midway through Phase I, students in Phase II were asked these questions before nearly every lesson (I inadvertently omitted questions about the pre-class preparation from the polls for lesson 5). Students self-reported both reading and watching the videos before class 102/132 (77%) times (the denominator here is the number of times any student in either course from Phase II completed the pre-class quiz). There were only five instances of any student coming to class on-time to take the quiz without having either read or watched the videos, and three of those were from the same student. Among the 25 instances when a student reported having done only one or the other of the pre-class activities, 22 (88%) were students who read the lecture without watching the videos. Thus, 94% of students who took the pre-class quiz had at least read the lesson text prior to class.

Ten students from Phase II provided feedback regarding their use of the videos to prepare for class. Of these ten, seven expressed uniformly positive evaluations of the videos. For example, Tobias wrote, "I found them useful, no need for any changes," and Dolly wrote, "Videos were helpful; don't change them! Nice to be given the lecture before class to free up time for specific questions and practice problems." Two of the remaining three students also described the videos as helpful, but suggested possible areas for improvement such as following along with more concepts directly from the readings or adding "a capstone video for the problems, question types, approaches, etc…citing in which lesson the different concepts can be found." Only one student, Fiona, expressed criticism: "I found them moderately helpful at the beginning but after a while they got repetitive." The overall increase in the distribution of positive and

constructive comments from Phase I to Phase II may be explained in part by my deliberate effort to avoid repeating content from the videos during class.

Another possibility is that some students in Phase II felt the repetition was helpful rather than unnecessary. This explanation is reinforced by Grace's comments during my interview with her. When I asked whether she felt the lectures in class unnecessarily repeated material from the videos, she said, "They were a little bit redundant, but it kind of helps to have a live person going through it. The same reason that I liked having videos with the lecture text is why I liked having you [lecture] after watching the video." For Grace, at least, the repetition of content in different modes—written, video, and live lecture—was perceived as beneficial.

Group work. Based on the feedback of students regarding cliques and my own observations of groups during Phase I, all groups for Phase II were randomly assigned for each lesson. As a result, Phase II students worked with a greater variety of their classmates, and the overall classroom environment seemed to improve. For instance, four of the female students who met each other in the second course formed a study group and, as one of them mentioned during an interview, had met socially on several occasions more than six months after the course ended. Also, during the final class, one student brought "lucky" LSAT pencils, and another brought homemade cupcakes for me and all of their classmates. In ten years of teaching previously, no student had given gifts to the rest of the class during the course, and only one student had ever given me a gift. This is suggestive of the possibility that these students felt a greater sense of classroom community than prior students had.

Although the random, repeated reassignment of group members seems to have had a positive effect on classroom atmosphere, some students felt that it had drawbacks, as well. Grace noted, "It might make students more comfortable if we didn't rearrange them [the groups] so often. It takes me a few questions to warm up to my new group members each class, so for me it would be beneficial to change groups less frequently." Franklin wrote, "If there were consistent groups, or regular rotation between groups, I think it would allow members to become more comfortable with one another and…help them develop the style of peer-teaching and learning that work best for the members of that group." Both students raised valid concerns and it is likely that using different groups for every class was not the optimal response to the problems that arose with students choosing their own groups during Phase I.

Themes from the discourse occurring in courses using a flipped curriculum. I recorded my own comments for each class during both phases, and the group discussions for three selected lessons (lessons 2 and 5 in both phases, along with lesson 6 in Phase I and lesson 7 in Phase II). These lessons were selected because I believed the complex concepts would allow me to observe both effective and ineffective group processes. After transcribing and repeatedly reading the discussions, I noticed a number of themes relating to the effectiveness of group work in both student-student and student-teacher interactions. The themes below are illustrated with thick descriptions, excerpts from the transcripts, and, where appropriate, counter-examples.

Theme 1: Complacency reduced group effectiveness. In certain cases, students assumed they understood more than they did, causing them to subvert the discussion

process. For example, one group consisting of Kyle, Alberto, Alex and Gonzalo in Phase I performed very well on the first practice game in lesson 5, with all four students getting all five questions correct. This group was justifiably pleased, and Kyle led a perfunctory discussion of a few wrong answer choices until I gave them permission to stop recording for that game.

Perhaps because of false confidence from game 1 or because I allowed them to stop recording discussion of that game, the remainder of the group's discussion during this lesson was disjointed and abbreviated. The other three groups in this lesson averaged 72, 72, and 127 statements per student, whereas these four students combined spoke only 116 statements. The total amount of recorded audio for this group during this lesson was only nine minutes, far less than any of the other groups. At one point, Alex asked Gonzalo, "You wanna talk?" and Gonzalo responded, "Not really" before shutting off the recorder. By this point in the course, Kyle had scored twenty points higher than either Alex or Gonzalo, and as will be shown later in this chapter, he was typically willing and able to share his understanding with his classmates. I believe that Alex and Gonzalo missed a significant learning opportunity by choosing to spend so little time on task during this lesson.

Also, the few conversations that occurred were primarily between Kyle and Alex, with Alberto and Gonzalo seldom contributing. Gonzalo's occasional comments reveal that he was following the discussion closely, whereas Alberto's interjections reflected little understanding of what the other group members were discussing. The following excerpt is taken from 4:09 to 6:49 of the group's discussion, and I have bolded Alberto's

questions about the group's progress (in the tables representing group discourse, each column represents comments from a different student and comments within the same box were spoken simultaneously or with overlap):

Alberto: **What did you say?**		
	Alex: I said that on question 10, J had to have at least the postcard. It's telling it cannot be a complete and accurate list of pieces of mail for J, that's how I got E for 10. What did y'all put?	
		Gonzalo: I put D, but I kinda just went with it. I was running out of time and I didn't really go through all the steps.
		Kyle: Yeah, I was process of elimination. I started at the top and worked my way through and E couldn't be possible because if you J with F, M, and S, then…
	Alex: If you go off the rules, I believe J has to have…	
Alberto: **What are we talking about?**		
	Alex: If R has L…	
Alberto: [*interrupts*] **Which number?**		
	Alex: [*crosstalk*] …then J must have P.	
		Kyle: [*crosstalk, to Alberto*] Ten. Right, exactly. There you go.
	Omitted turns for brevity	
	Alex: It did work, yeah.	
		Kyle: So it's not the right answer. It's a cannot question.
Alberto: [*interrupts*] **What did you get for 11? You got B?**		
	Alex: [*crosstalk*] Yeah, cannot work. It was the survey, right, that didn't fit with R?	
		Kyle: [*crosstalk, to Alberto*] I got B. [*to Alex*] It was the letter, the survey…
	Alex: If L is R, then P is J.	
		Kyle: You know, which one ever has…it could have been G, G, F.

	Alex: I believe it was the survey that didn't fit with R. There's no really rule for…uhhhh….
	Gonzalo: G, that works, right?
	Kyle: Are we wrong about that?
	Alex: I don't know.
Alberto: Yeah, it's the letter and the survey.	
	Alex: What's that?
Alberto: **You guys are talking about which number?**	
	Alex: 11.
Alberto: Yeah, it's B.	
	Kyle: Right.

Over a two and a half minute period, Alberto asked five clarification questions, two of which are requests to repeat information other group members had recently spoken. Alberto's inability or unwillingness to keep up with the group not only caused unnecessary repetition, it also resulted in interruptions and crosstalk. At one point, Kyle carried on conversations with both Alberto and Alex simultaneously because Alberto missed Kyle's comment from fifteen seconds previous. And, despite Kyle's efforts to multi-task, Alberto's question about the answer to number 10 went unanswered. When Alberto finally caught up to the rest of the group, he quickly resolved their confusion about question 11, which suggests that he may have been able to offer other contributions earlier in the discussion if he had been keeping pace with the others. Following their perfect performance on the first game in this lesson, these students put less effort into engaging with each other and the material than was typical for them in other lessons.

Another pair of students from Phase I who evinced signs of overconfidence and complacency was Barry and Katrina. Katrina's initial practice test score (157) was the

second highest in the course and although Barry's score was at the median (145), he was among the most vocal and confident students from the beginning. During lesson five, Barry and Katrina convinced each other that I had improperly drawn an inference during one of the games. Below is an excerpt of their discussion:

Katrina: But H and P would both have to be in O.
Barry: Yeah, you're right. I didn't feel like he was right when he said that P's not in O…
Katrina: Me, neither.
Barry: I was like, I really, really had doubts and then I just wrote it off, almost suppressed.
Katrina: Hey! [*to me, laughing*] We think you're wrong.
Me: Ok.
Barry: [*to recorder*] Please take note.

When I explained what they had misunderstood, Barry expressed chagrin and went so far as to apologize into the microphone after I had moved on to a different group. A few moments later, however, Katrina insisted, "It was how he said it that made me think they both had to be together" to which Barry replied, "Yeah, I specifically remembered that as well." Although both students had accepted my explanation, they remained surprised at being wrong and convinced themselves that the misunderstanding was at least partially due to a mistake they believed I had made.

On at least one other occasion, Barry and Katrina expressed similar surprise at being incorrect. During lesson 8, they had finished a logical reasoning question and both had picked the same answer. Rather than discussing the argument or their thought process for selecting that answer, they initiated an off-topic conversation. When I indicated that

they had both chosen the same wrong answer, both students gasped in disbelief and started to rework the problem. This instance stood out to Katrina, who recalled it during our interview several weeks after the course had ended.

Both students were regularly among the first in the class to finish discussing problems, and spent much of the group time flirting with each other instead of solving practice problems. In our interview, Katrina noted that her classmates "thought [she] was annoying because [she] was always talking to Barry." Several times when Barry and Katrina were grouped with other students, the recordings indicate that these two focused primarily on each other and did not attend to the comments of their groupmates. During the same lesson as the above excerpts, Barry and Katrina finished discussing a game before their groupmates, Cindy and Bonnie. Cindy asked both Barry and Katrina about their answers to two questions from the game and they briefly replied before moving on to discuss Katrina's new diet. It is clear from Cindy's tone that she wanted to discuss the questions further and during the next two minutes, she asked about the same questions two more times. However, Bonnie did not know the answer, and Katrina appeared to be too busy discussing various types of seafood with Barry to explain how she had solved the questions. Although both Barry and Katrina attended the majority of lessons, neither student completed tests 3 or 4.

By contrast with those students whose complacency resulted in shallow engagement with the course materials, the most effective students sought not only to get the right answer using the right methods, but also to understand if they would have been able to derive those methods without my assistance. In this example from lesson 5 of

Phase I, Andrea, Kyle, and Colin discussed a logic game for which there are two possible appropriate approaches:

Andrea: Let me ask you: how did you solve that? Did you infer that by looking or did you just start testing combinations?	
	Kyle: Brute force. I went down the list, first one worked, second one I got stuck, so…
	Colin: I mean, it depends. Like some of these, I guess if you can, quickly go through and pick that. If not…I mean, sometimes you can get inferences from the previous problems, too.
Andrea: Yeah.	
	Colin: I mean, not always, but…
Andrea: Where I'm running into problems, is that basically, I'm solving a lot of these just by trying a bunch of combinations, and that seems like a really time-consuming way to go about it.	
	Colin: Well, with that diagram, that one up there, it's a little harder to do testing it.
Andrea: Getting L/P.	
	Colin: I mean, like I said, for that diagram, I can see why you would really need those inferences to get through it.

Colin and Andrea had chosen different approaches for the diagram to this game, and Andrea wanted to know how the others had approached a specific question. Her question led to a general discussion of the advantages for each diagram and was part of a pattern within this group of discussing more than just the correct answer to each question. The group's fourth member, Bertha, was an active participant in discussing specific questions, but was generally silent when the group began to analyze their processes.

Later in this lesson, Andrea, Kyle, and Colin continued to discuss an inference from the same game well after they had already finished discussing all of the questions.

Andrea: Yeah, so…yeah, I guess maybe I would like to ask him again how he got to the point where J always has to have either L or P, because I used that answer to quickly answer 9 and 10 without doing any further work.	
	Kyle: Um-hm. I did the same.
Andrea: Yeah. So that I'm having trouble…	
	Kyle: It because it appears twice in the rules…one time is the end of the sequence, not PJ —> not RL. Then it appears in the first rule L not with G. So from that, he pulled those together and saw that it was twice.
Andrea: So that is how you get to…	
	Colin: He shows you right there how he got it. If L is in R, P is in J. If P is not in J, L cannot be in R. If L cannot be in R, L has to be in J because L can't be in G. So L would have to be in J, so L or P are in J.
Andrea: So we can derive the L P rule through the placement of L? So if L is here, P has to be there. If P is not there, that means….I see. I mean, I understand now that I'm looking at his diagram, but I really don't feel like that's something I would have come up with on my own.	
	Kyle: Um-hm.
	Colin: It helps that there are only three possibilities. In that diagram, you can only have G, J, or R.
Andrea: Right.	

This discussion was well-balanced and evinced careful consideration of both the problem set and the students' own thought processes. Although all three understood the inference and had properly applied it, Andrea wanted to explore how the inference was originally derived and suggested asking me to explain it again. However, Kyle and Colin were able to explain it well enough that she did not feel the need to ask me about it, although she remained concerned about her ability to find it on her own. Colin attempted to assuage Andrea's concerns by saying, "You would have still come to that conclusion on your own because it bases off that rule that R is L then J is P. So even if you didn't get that

inference right away, when you were doing the problems, you would've seen it." The depth of processing and the empathic responses during this exchange are both noteworthy. In her survey responses, Andrea noted that "explaining concepts to others was an effective way to learn," which is well-supported by the exchanges above.

Indeed, Andrea was among the students most gifted at articulating her own thought process, regardless of the group in which she found herself. During lesson 6, for example, Andrea explained her approach on the first question for game 1 as follows:

> Um, so the way I started it is I wanted to pick the easiest rules, so I first started with the three member rule, because that was visually easy to eliminate. So I started with that one. And then from there, I did the M rule. So that knocked out A. And then...I think I started with the conditionals, like the F not K, that got one out. K and J together, that got another one out, and by that point, I was left with B.

Here, Andrea explained not only why the correct answer worked, but also the order in which she eliminated each wrong answer and her rationale for prioritizing various rules. She was highly conscious of her own thought process and able to defend each decision clearly. She added, "I sometime try to steer conversation to why wrong answers are incorrect or discuss the methadology [sic] we took to get there." I believe that this process was not only beneficial to her, but also to her groupmates. Colin agreed, saying, "I like the group work. I think explaining and listening to other people explain really does expedite the learning process." Even after answering questions correctly, the students in this group showed few signs of complacency and regularly evaluated the usefulness and appropriateness of various problem-solving strategies.

Theme 2: Group work allowed students to reinforce new concepts with each other. In a traditional course, it is often unclear whether students are simply guessing

correctly or if they have learned to apply the recommended processes to confirm correct answers. However, in functional groups, the discussions often compelled students to articulate previously unclear thought processes and allowed them to reinforce the application of recommended approaches. Below, I have indicated several instances of students correcting each other's approaches or seeking input on their own approaches.

For example, my typical approach in each course (whether using a traditional or flipped curriculum) was to teach students that, rather than spending a significant amount of time trying to evaluate confusing answer choices, they should designate these answers as "contenders" and move on to consider other answer choices. Many students resist this suggestion, preferring to determine conclusively whether an answer should be selected or eliminated before proceeding to the next answer. Although thorough, such an approach is often unnecessary or inefficient, especially when students encounter a confusing answer at some time prior to reading the correct answer.

In some of the groups, the students reinforced each other's use of the recommended approach. For example, in lesson 2 of Phase II, Nate, Debby, and Ernesto were discussing a logical reasoning problem. All three agreed that answer choice A was incorrect, and that B seemed plausible. After unsuccessfully attempting to deconstruct answer choice C, Debby said, "Ok, let's leave that one, because it's obviously a contender if we can't figure it out." Nate and Ernesto agreed, and they were able to eliminate quickly answer choices D and E. In the excerpt below, Nate and Debby discussed the remaining two possible answer choices:

Nate: Ok, so B and C are the contenders.	
	Debby: B and C.
Nate: I think the answer's B, but C could be…I just don't really feel like I can explain it. [pause]	
	Debby: Ok, because the question is offering, like, could be true, right? Which is what B does, but C says must be true.
Nate: Hmmm, I didn't understand that.	
	Debby: Because we know that Jean definitely hasn't had jaundice, but Mary could have sinusitis.
Nate: Yeah, I did notice that.	
	Debby: So it is B.
Nate: You're saying Mary…could you say that again?	
	Debby: I'm saying that Mary could have sinusitis and that's why she lost her sense of smell, which is the phrasing in the stimulus, but in C, we know for sure that Jean hasn't had jaundice, because Jean is a blood donor.
Nate: Ok.	

This excerpt demonstrates several effective group behaviors. First, Nate recapped which answers were still under consideration. He then expressed an inarticulable preference for B, which is the correct answer. In a traditional setting, Nate may have stopped there, satisfied that he had done enough to work out the question. But in this group, he and Debby addressed the question further until both were fully satisfied. Of particular note is the manner in which Nate prompted Debby for further explanation and the increasingly specific nature of her replies. It is possible that neither student would have understood the question as well without the other's input, even though both of them would have likely answered it correctly (it is also noteworthy, although less encouraging, that Ernesto did not speak for nearly ten minutes during this section of the recording).

At the same time, some students remained reluctant to apply recommended approaches fully. From the same lesson as above, Nate, Debby, and Ernesto discussed the next question. Here, the group once again followed the recommended process, but both Debby and Ernesto expressed reservations about doing so.

Nate: [*reads answer choice B*]	
	Debby: I mean, that is…it could be true, but it's not something you can infer from the passage. [*pause*] Like there's not enough information to say that that's…
Nate: Really?	
	Debby: I don't think so.
Nate: What is the part of the information in B that's not in the stimulus?	
	Debby: That they do not help patients before they reach the marketplace.
Nate: Ok, I guess I saw that as negation. I saw the word before as negation, because here it says after the transfer.	
	Debby: Oh, I see.
Nate: So, I saw before as the negation of after. [*pause*] So we looked at that one differently.	
	Debby: But it's still not right.
Nate: I think…why isn't that right?	
	Ernesto: Ooooh, I hate this.
Nate: Jason [me] wants us to ask why things are not right. Jason says the origin of wisdom is knowing why things are not right.	
	Debby: I don't have time to learn why things aren't right. I just want to learn why they are right.
Nate: So…	
	Debby: Ok, because it says only after. Only means that like that's the only scenario in which they help patients so they can't possibly help them before.

Nate: So it says here do not help patients. So you're saying can't possibly help before and it says that before new therapeutic agents reach the market, they do not help patients.
Debby: Ok, I see what you're saying.

Although Nate described my instructions with an air of mock importance, it is clear from his later statements that he was intent on thoroughly analyzing this answer choice. When Debby claimed that she didn't have time "to learn why things aren't right," Nate pointedly ignored her comment and instead prompted Debby to continue explaining her rationale. By encouraging Debby to explain why she thought B was wrong, Nate was able to understand her thought process and help her see why it was actually correct. Interestingly, neither Nate nor Debby knew the correct answer at this point, but completing the process, however reluctantly it may have been for Debby, helped both students to clarify the problem.

At other times, a student's attempt to reinforce a recommended approach failed for idiosyncratic reasons. For example, in lesson 5 of Phase I, Katrina, Barry, Cindy, and Bonnie were discussing the final question of a relatively difficult game. By this point, Katrina had shared with the group that she had achieved a perfect game, which resulted in her becoming the de facto authority for the rest of the explanation. In the following exchange, Bonnie asked Katrina to explain her approach, while Cindy explained why she had not chosen to follow that method:

Bonnie: How did you do 12?
Katrina: Well…[*professorially*] [*Barry laughing*] [*Katrina reads question*]

	Barry: Oh my God.
Katrina: So, it couldn't be…yeah, so, really for that one, you just have to….	
	Bonnie: Plug in?
Katrina: Yeah, I got to the second one and tried [them].	
	Cindy: I can't even do, like, I gave up on it.
Barry: [*crosstalk*] It sucks.	Bonnie: [*crosstalk*] Ok, the M and the S are together.
Katrina: Yeah, but once you get going…	
	Cindy: See, I don't put the stuff because I want to [*inaudible*] drawing because I want to not know the answers.
Katrina: You what?	
	Cindy: [*crosstalk*] I don't try to put the right answers in so that way I wouldn't know so I can reuse my book, so that's why, when he wrote it down, I'm like, what?
	Bonnie: [*crosstalk*] Ok, M can't be first, so M and S, so then that means it has to be here.
Katrina: Oh, ok.	

Here, Katrina has demonstrable credibility among her group members (Barry's initial reaction was in jest and he later congratulates her for doing so well on this game). As Katrina explained her approach, Bonnie attempted to rework the problem. However, Barry and Cindy lamented how difficult the concept was, and, to Katrina's surprise, Cindy explained that she preferred not to even attempt the recommended approach because she wanted to be able to reuse her book (I surmise that she was concerned that having the answers written into the book would preclude her from retaking the questions on her own at a later time). Although there are a number of possible solutions to Cindy's

preference that would allow her to record her work in class and use the book again later, such as using pencil or writing out the work on a separate sheet of paper, her insistence on not writing down hypothetical solutions prevented Katrina from successfully reinforcing the recommended approach.

Theme 3: Group discussion allowed members to identify and correct mistakes. I also observed instances of students defending incorrect answers or assumptions to their groupmates, only to realize in their explanations that they had made a mistake. During lesson 6 in course 1, Kyle and Andrea discussed their initial inferences for game 1. Kyle listed three inferences: F and K could not be together, F and J could not be together, and K and J must be together. In fact, only the first inference is actually valid given the rules of this game.

Andrea asked Kyle to explain each inference and they jointly determined that the first inference was accurate. Next, Kyle explained that he had discovered during the game that his third inference—K and J must be together—was incorrect. He said, "As the questions got in, I realized this was wrong. K and J didn't have to be together; you could have J without K." Here, Andrea started to redirect the discussion to the second inference, saying, "Wait, so wait. J not F, but F..." As Kyle was directed to reconsider this inference, he realized that he might have misunderstood the rule: "Uh, you could have J and F together...So I think you could have a situation where F is out...err...K is out, and F and J are together."

Andrea hesitantly restated Kyle's new conclusion, before Kyle repeated it again. This led to the following exchange:

Andrea: Ok, yeah.

Kyle: Is that right?

Andrea: Umm....

Kyle: Because K wouldn't be on a committee so that takes that rule out.

Andrea: Ahhh!

Kyle: So if K is on a committee, J is also there. If K is not on a committee, you don't know.

Andrea: Right.

Thus, in just a few minutes, Kyle made a faulty claim, realized his mistake, drew the proper inference, and convincingly explained it to his group, prompted mainly by Andrea's subtle insistence on understanding his rationale. In fact, throughout the exchange, Andrea served primarily to reflect back Kyle's reasoning to him. I find this to be a compelling example of co-construction of knowledge because neither student had fully understood the issue before they began discussing it and both had firmly mastered it by the end of their conversation (there was no audible indication of understanding from Clarissa, but she did not seek additional clarification, either, and later demonstrated the correct application of this inference).

In other instances, however, group members compounded or reinforced each other's mistakes. In the following example from lesson 2 of Phase II, Dolly, Ernesto, and Tobias were attempting to solve a logical reasoning problem with conditional reasoning. The specific task was to match the pattern of reasoning in the stimulus to that in one of the answers. The correct answer was B, which used the same pattern of flawed reasoning as the argument, but this group struggled to identify the salient features of each answer that would have allowed them to identify similar arguments:

Dolly: Ok. So let's diagram A. [*pause*] I feel like this is a really awkward club here. Alright
Tobias: So if fair-skinned…
Dolly: So if you're fair-skinned, then you suffer from sunburn. So, if FS, then SB.
Tobias: Suffer from sunburn...
Dolly: So…
Tobias: So if you don't suffer from sunburn, then you're not fair-skinned.
Dolly: But I guess it's saying Gaston is FS and suffers from SB. Ok, that seems like it matches. So… [*prolonged pause*]

Throughout this exchange, Dolly sought input from the other group members and was reluctant to make conclusive judgments on her own. However, her tentative, but incorrect, conclusion that A "seems like it matches" passed unremarked, and Dolly decided to move on to the next answer.

Dolly: Ok, let's go to the next one and diagram the next one [*sheepishly*] People who suffer from sinusitis, lose their sense of smell. So if you suffer from sinusitis, then you lose…
Tobias: Wouldn't that…? Wouldn't that losing sense of smell be the sufficient condition?
Dolly: I don't know. I'm really bad at this. I did half the homework and still could not answer them. [*Dolly and Tobias laughing*]
Ernesto: I just don't know how to diagram it, but it's wrong, isn't it?
Tobias: I think it's wrong.
Dolly: Ok, we're all…good job, team.

This is the correct answer, and Dolly correctly started by placing "suffer from sinusitis"

as the sufficient condition. However, Tobias hesitantly suggested that "sense of smell" might instead be the sufficient condition, upon which Dolly expressed doubt in her own ability. Dolly's claim that she had done half of the homework is significant because homework problems were not assigned prior to class. Even though she struggled to complete the problems, it is likely that she had practiced more questions than either of her groupmates. It is clear, however, that this was insufficient for her to be confident in her approach when it was questioned by another group member whose approach was incorrect.

Without diagramming answer B, Ernesto's sense that the answer is wrong is technically accurate but misleading, since the statement is flawed but the correct answer is supposed to be flawed for this question. Tobias quickly agreed that B was wrong without any discussion of how it failed to match the argument, and Dolly assented. At this point, it is clear that the group members misunderstood either the initial argument or the task they were expected to complete.

Dolly: C. People who have suffered from jaundice cannot become blood donors. Jean has suffered from jaundice, so Jean cannot become a blood donor.
Tobias: Isn't that a loser? Loser.
Dolly: Right.
Tobias: Loser for sure.

Answer choice C is eliminated quickly, but again without discussion.

Dolly: People who are colorblind cannot become airline pilots. Arthur is colorblind, so Arthur cannot become an airline pilot. [*Tobias laughing*] Ok, so that's…
Tobias: Arthur is CB. I think it's a contender.

Dolly: Alright…
Ernesto: It looks decent to me.
Dolly: Alright. [*laughing*]

Answer choice D is deemed a contender, probably for the same reasons that the group liked A, since both answers follow the same pattern of reasoning. However, this rationale is only inferred because the group members again failed to describe their reasoning.

Dolly: People who are diabetic cannot eat large amounts of sugar. Frieda is diabetic, so Frieda is on a special diet.
Tobias: That has nothing to do with it.
Dolly: Right. So we have two contenders: A and D.
Ernesto: I would guess D.
Dolly: Yeah, I'm gonna say D also.
Tobias: Because of what…the cannot?
Dolly: [*crosstalk*] Yeah. Ernesto: [*crosstalk*] Yeah.

Answer choice E is also eliminated, leaving the group with two wrong answers to consider and no clear criteria for making a distinction. Tobias asked why D is better than A, and noted that the word "cannot" is present in both the argument and this answer choice. However, "cannot" reveals nothing about the pattern of argumentation and is irrelevant in determining the correct answer. Dolly and Ernesto's immediate agreement is not based on principles from the lesson material or video, and appears to be an ad hoc justification of an instinctive decision. The group's overall inability to deconstruct the arguments facilitated multiple errors and was observed in several other groups, as well.

Theme 4: Lack of group cohesion reduced group efficiency. Throughout the study, individual students often worked faster or slower than the other group members, resulting in less cohesive groups. This was noted by several students when discussing obstacles to group work. In the mid-course survey, Chris wrote, "Group members tend to explain their setups and answer too quickly for me to get a firm grasp on the concepts," and Andrea added, "Sometimes other group members weren't keeping up with lecture/lesson materials and didn't necessarily work at the same pace." Franklin wrote, "[T]he work pace of the individual varies and at times makes it difficult for everybody to be at [sic] similar place with a generally equal understanding."

Mismatching of pace seemed often due to one student's inability to work as quickly as the other group members, but also seemed to result from impatience, which led to broken turns, disorganization, and confusion. For example, during lesson 5 of Phase I, the students were asked to discuss their solutions to a logic game involving a fruit stand. The game has six questions, and the rules involve which combinations of fruits can appear concurrently. The rule that is most confusing for many students states: "If kiwis are not on sale, then tangerines must be." I am usually compelled to explain the implications of this rule several times before students understand that either kiwis or tangerines, at minimum, must always be on sale. Correctly answering two of the six questions depends on understanding this implication, in particular, question two which asks which of the fruits could be on sale by itself. Given the above rule, the only possible correct answers would be either kiwis (not listed as an answer) or tangerines (answer choice D), but several students have difficulty understanding why figs (answer choice A)

cannot be on sale by itself. Although all of the group discussions from Phase 1 reflected difficulty understanding this game, one group in particular serves as a noteworthy example of how an ill-conceived (or ill-understood) group process can make a hard problem even more challenging.

This group started off by determining how far everyone had come in solving the problem set. Gonzalo and Alex had completed the first three questions, and Alberto had only completed two questions. Colin had finished all six. When solving logic games, the other group members often deferred to Colin, who was generally the fastest and most successful in the group at completing this section. In this case, however, Colin was so far ahead that he chose not to participate while the others discussed the first two questions and he missed the fact that he had not understood this portion of the game as well as Gonzalo had.

While Colin was looking over questions four, five, and six by himself, Gonzalo explained question two to Alex, particularly how he was able to eliminate figs and select tangerines, and then quickly covered question three. He then re-explained his rationale for number two, apparently to Alberto, saying that "if the stand does not carry tangerines, then it carries kiwis. So if you picked figs and that's the only one you got, then you gotta carry kiwis, too, and it [figs] can't be alone [since kiwis would be added to the group, as well]." Here, Gonzalo's explanation is precisely correct and he is confident enough to explain it to two other group members. Critically, however, Colin missed the explanation, which would lead to significant confusion and misunderstanding.

At this point, several seconds passed with no discussion. After I encouraged the

group to continue discussing, Colin picked up the discussion with question four. He explained his rationale correctly, and then sought confirmation from the other group members:

> I thought 4 was C, because if not W, it can't be O. And then if you leave out W and O, you've got PTKF, but you can't have K and P together, so at most it could be three. Does that make sense? Did I do that right?

However, none of his teammates had completed number 4, so his question was met with silence. For the second time in just three minutes, several seconds of silence passed, and no one asked a question or ventured an explanation.

At this point, Colin returned to question two, and Gonzalo began to explain his reasoning for a third time. However, before Gonzalo could explain the correct reasoning, Colin asked why pears couldn't be present alone. Gonzalo was clearly caught off-guard and failed to realize that his previously explained rationale for eliminating figs alone would also suffice to explain why pears alone was wrong:

> Gonzalo: I'll just tell you the reason for why I didn't pick A and then, I don't know, we'll have to review that one. The reason I didn't pick A—because I got it down to A and D—the reason why I didn't pick A is because let's say you were to pick A, figs, right?
>
> Colin: Um-hm.
>
> Gonzalo: That's the only one you're going to sell. Look at rule number two- the stand does not carry tangerines, then it carries kiwis. Ok, well then if you only have figs…
>
> Colin: Yeah, I see what you mean.
>
> Gonzalo: You don't have tangerines, so that means you have to carry kiwis, so that's two.
>
> Colin: Yeah.
>
> Gonzalo: So obviously it can't be. But it's like you have to really…you know what I mean?
>
> Colin: Yeah.

Gonzalo: Because it doesn't say specifically, like, oh if you have figs…

Colin: So you picked D, right?

Gonzalo: I picked D, but I mean, it could have been just because my thing was wrong, and it told me not to pick C.

The discussion then turned for several minutes to Colin explaining to Gonzalo why pears could be correct, and Gonzalo realizing that he had a mistaken inference about a different rule. However, Gonzalo still failed to realize that his process for picking tangerines over figs would answer Colin's question about picking tangerines over pears.

Eventually, Gonzalo did explain his rationale to Colin, but was considerably less confident about his thought process: “But you know on the real exam I could have been completely wrong." Colin then explained that he wanted to find out why tangerines couldn't work, apparently unconvinced that tangerines alone is the correct answer. A minute and a half later, Colin said, “It is tangerines,” and using a different rationale from Gonzalo's, explained that pears could not be alone because they would have to include tangerines as well. Colin concluded, “So, I think you got the right answer, just the wrong way.” Colin’s rationale for eliminating pears was accurate, but he was incorrect in dismissing Gonzalo’s approach and never realized that Gonzalo’s approach was simpler. It seems plausible, at least, that if Colin had listened while Gonzalo explained number two to Alex and Alberto, he might have realized that Gonzalo’s approach was the most effective.

After spending several minutes on question two, Colin and Alberto briefly discussed question five, before Colin discussed question six with me. A few minutes later, Colin asked the others what they chose for question four. In order, then, this group

had discussed question two, three, two, four, two, five, six, and four. Furthermore, question three was barely discussed (Gonzalo explained it away very quickly) and although Colin explained four quite well, no one else seemed ready to discuss it at that point. The transitions between questions were abrupt, with little or no closure for previous questions and were interspersed with periods of silence ranging from five to ten seconds long. In all, the conversation was halting, disjointed, and painfully awkward, with a long tangent that obscured one of the key points for resolving the game. Although it is probably incorrect to suggest that Colin's pace caused all of the group's dysfunction in addressing this game, it is instructive to note that when Colin set a pace that the other group members could follow in the next game, the discussion improved significantly.

Theme 5: Group work revealed student misconceptions to the instructor. There were several instances when the group work process allowed me to understand my students' misconceptions more easily than the traditional course format had previously allowed. One such instance occurred during the second lesson of Course 2. This lesson introduces students to conditional reasoning problems, involving the identification and manipulation of necessary and sufficient conditions in logical reasoning arguments. Generally speaking, there are two approaches to diagramming the statements in a logical reasoning problem: students can either use acronyms based on particular phrases from the stimulus (i.e., students might represent the condition "Gerald is red-green color blind" as "$RGCB_G$") or use generic letters to abstractly represent the conditions (i.e., "Gerald is red-green color blind" would be "A_g", and the following condition would be "B_G").

In this lesson, I first demonstrated the acronym approach, but I noticed that one

student, Fay, was using the abstract method. Fay explained the process to her groupmates, both of whom seemed to understand it quite easily. Her group arrived at the correct answer more quickly than either of the other groups and had little trouble applying the abstract process. As a result, I chose to explain the abstract method to the other students earlier than usual, and I praised Fay for her explanation to her classmates.

As I listened to the small groups apply this approach, I noticed that some students were having difficulty remembering which letters represented which conditions, but they were able to answer the questions correctly regardless. However, as the groups discussed question 6, it became apparent that using abstract representation was generally counterproductive. In the transcript below, Dolly tried to encourage Don to explain the diagramming process for the group:

> Dolly: So, you've got yours labeled nicely, so share and like...
>
> Don: Well, I don't know why, but inspired musical performances is A.

As Dolly realized that Don was either unwilling or unable to articulate the process properly, she took over for him, explaining the next steps to Don and Tobias:

> Dolly: Ok, so that's A, because of if, which makes it the necessary, right?
>
> And then the audience will be treated to a good show, so B. And then you come down here so you have unless which automatically puts it here [necessary] and then you negate not good show, right? Here, so that's B, so you can move C to here. Does that make sense?

Although Dolly explained the modifiers correctly and accurately represented all three conditions, both Don and Tobias remained confused. When I heard their discussion, I attempted to explain the abstract method again to their group, and soon realized that even Dolly had difficulty interpreting her own representations:

Me: There's nothing you could ever do that infers the necessary condition is absent. So anything that's at the end of your arrows, you will never correctly infer is gone. It can't be done. They don't work that way because of what kind of relationship exists, so whenever it says get rid of D, that's wrong.
Dolly: Ok.
Me: How could you infer that B doesn't happen?
Tobias: So, no "sophisticated listeners"?
Me: Yeah, say I want to prove there are no "sophisticated listeners." How would I do that?
Dolly: Can you get rid of A?
Me: Nope, because that goes the wrong direction. That goes down. I can't say no A, no B.
Dolly: So, no C instead.
Me: No C…
Dolly: …that makes sense.
Me: If I get rid of C, I'm going to be able to get rid of B.
Dolly: Ok.
Me: If I get rid of D, I'm going to be able to get rid of B. Right?
Dolly: Right.

At the beginning of the discussion, I used the abstract conditional representations of A, B, and C to reinforce what I had heard Dolly explaining to the others. However, as the discussion progressed, I saw that even though Dolly had previously led the discussion, she was confused about whether A or C was required for B. At this point, I abandoned the abstract representations in favor of the original terms:

Me: If wanted to show no "sophisticated listeners"…
Dolly: [*crosstalk*] It has be in front of it.
[*crosstalk*] I have to find something it depends on.
Exactly. And get rid of that. It's what it needs. Take away what it needs and then it falls… So basically anything that leads to something else, these two can't be removed. But if I take away the thing that it leads to, then I can say that thing didn't happen.

	Tobias: "Sophisticated listeners" depends on "good show."
Me: Depends on…does "sophisticated listeners" depend on "good show"? Is that right?	
	Dolly: I don't know.
	Tobias: I thought we were going…
	Dolly: [*looking at question*] "Musical roots."
	Don: Yeah, there will not be "good show" unless there are "sophisticated listeners."
Me: So actually "sophisticated listeners" depends on "musical roots." That's the only thing it depends on. "Sophisticated listeners" is actually C, so "sophisticated listeners" only depends on "musical roots." That's the only thing it depends on, right?	
	Dolly: Right.
Me: If it's present, there have to be "musical roots." The only way to get rid of "sophisticated listeners" would be to get rid of…	
	Dolly: [*simultaneously*] "Musical roots." Tobias: [*simultaneously*] "Musical roots."

Even with the concrete terms, Dolly was unable to recall what "sophisticated listeners" depended on (although she did quickly realize that it was "musical roots" rather than "good show", which neither Tobias nor Don understood properly). However, by the end of the discussion, both Dolly and Tobias seemed to understand the proper relationship between the conditions, as they jointly completed my sentence. When I began to use the concrete signifiers from the stimulus, I noticed an immediate increase in the group's ability to discuss the problem (even if full comprehension required further instruction). After listening to members of the other groups struggle with the abstract concept, I encouraged the class to revert to using concrete acronyms.

In a traditional classroom, the fact that most students answered the question correctly, combined with my praising the abstract approach, may have obscured how many students struggled to apply this approach. At best, a confused student might ask for further explanation, but it is unlikely that I would have been able to determine the extent of the confusion. In courses using a flipped curriculum, I was able to overhear several different groups struggle and immediately revise my instruction.

Even more usefully, this episode prompted me to reconsider the value of both approaches. I have taught the course nearly fifty times over the past ten years, and the order of presentation of the major concepts has been essentially the same throughout that time. The course material emphasizes concrete representation in lesson 2 when conditional reasoning is first introduced, and then uses abstract representation in lesson 8. My tacit understanding of this sequence was that abstract representation builds upon and improves on concrete representation, because it subordinates the specific concepts to the underlying logical relationships. I also believed it to be more efficient because the diagrams typically involve fewer letters. Further reinforcing this belief was the fact that only a few students could understand abstract representations during lesson 2, and a significant majority of students could do so by the time they reached lesson 8.

When traditional students had struggled during lesson 2 to apply abstract representation, I had previously assumed it was due to unfamiliarity and lack of exposure to conditional reasoning rather than misalignment with the domain or content of the lesson. During this course, however, I noticed that even the best prepared students were far more comfortable discussing questions from lesson 2 concretely rather than abstractly.

As I listened to their discussions, I noticed that for stimuli with a single argument, the abstract approach was actually too simple as it tended to diminish the differences between various descriptions of conditional relationships. Although this simplification was beneficial for the kinds of questions addressed in lesson 8 (which contain arguments in the stimulus and each of the answer choices), it was detrimental for those in lesson 2 (which have only one argument). My desire for these kinds of insights into my students' learning was one of my primary reasons for attempting to implement a flipped curriculum, and it was rewarding to see such results, at least on a few occasions.

Theme 6: Mismatches between student-student interaction and student-teacher interactions interfered with learning processes. As I was circulating among groups, it was often difficult to ascertain what each group had been discussing before I came around to work with them. Asking the students diagnostic questions yielded mixed results in terms of both the thoroughness and accuracy of students' description of their progress so far. Also, although I frequently listened to portions of the students' discussion before joining in, what I heard was not always representative of the students' overall understanding. On several occasions, I did not realize my misunderstanding until I analyzed the taped discussions, days or even weeks later.

For example, in lesson 2 of Course 2, one of the small groups, consisting of Nate, Ernesto, and Debby, had been discussing a logical reasoning problem. Their task was to match an argument from the stimulus to one of the answer choices. Nate was leading the discussion and his analysis for each answer choice had been quite thorough. A few moments before I reached the group, Nate dismissed answer choice (C), saying, "It's not

saying Jean is not a blood donor; it's saying Jean is a blood donor. So it seems to not mirror the first one in the way that it is not negating the necessary condition.” Here, Nate successfully translated concrete terms from the problem into generic terms previously used during our class discussion.

When I joined the conversation, I saw that the group has already finished discussing three answer choices and I asked them to revisit answer choice C and describe the process it follows. Rather than answering my question, Nate said, “I don't really care. I just care that it doesn't...” Here, I interrupted and adamantly insisted that Nate (and the rest of the group) should in fact care about the wrong answers. What I did not discover until listening to the tapes was that Nate had, in fact, been doing exactly what I hoped he would do with the wrong answer choices. I had misunderstood his claim of not caring what answer choice C did as a more general dismissal of the incorrect answer choices rather than as the beginning of his explanation of why the answer was incorrect. Further, my interaction here was not only unnecessary, but possibly harmful, as it caused the group to doubt a process that had otherwise been both accurate and efficient.

Another instance in which my well-intentioned interaction with a group went awry was during lesson 5 of Course 2. Here, Harvey, Earl, and Tobias had been discussing a logic game in which students were asked to determine how each of five pieces of mail might have been addressed to three different housemates. The first question listed five possible distributions and asked which one of those could have been both complete and accurate. Harvey asked Tobias what he had done with answer choice A, and Tobias pointed out that it only had four pieces of mail. Harvey seemed to accept

this initially but then said, “But...you don't have to use all five.” Rather than referring back to the stimulus, which states that all five pieces must indeed be used, Tobias and Earl seemed to accept Harvey's interpretation and simply moved on to the next answer.

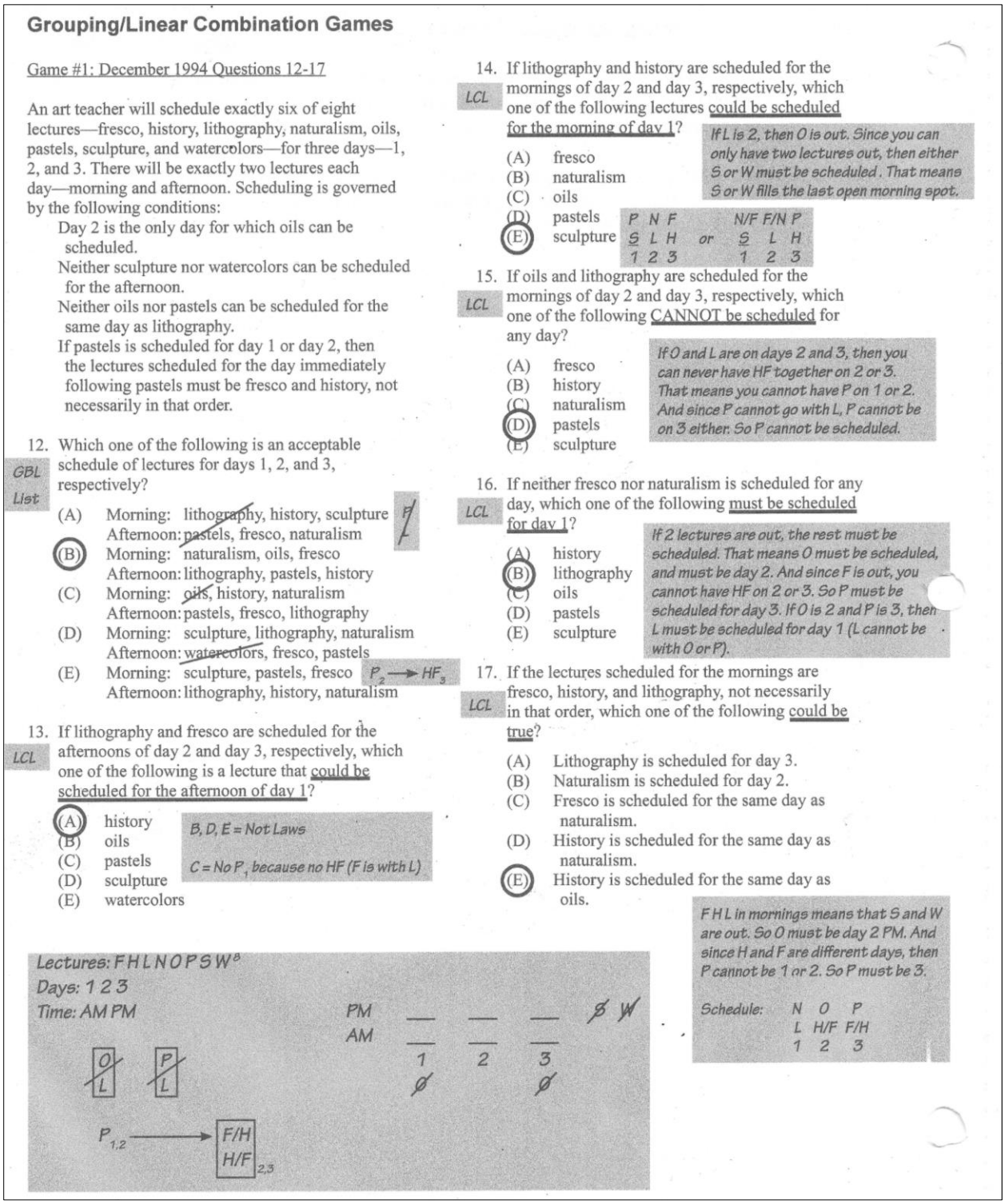
Unaware of this misconception, I overheard the group discussing another rule and briefly explained the proper interpretation. Over the next few minutes, this group struggled with this question while the other two groups had moved on to the next question. Harvey and Earl both agreed that B (the correct answer) seemed to be a good choice, but Tobias asked, “What's wrong with A?” While Harvey and Tobias both tried to identify a flaw in answer choice A (neither of them realizing that the only mistake in A is having four pieces of mail instead of all five), Earl eavesdropped on my conversation with another group. When I returned to his group, he immediately asked me to explain what I had told the other group. However, this concept pertained to question two, and I did not realize his group was still working on the first question. My explanation caused the group to reevaluate their answers for the first question, but did not help them eliminate answer choice A.

By now, eight minutes had passed, which was nearly the full time allotted for students to complete an entire game, and this group had not yet completed the first question. Tobias had been silent for more than three minutes, and Earl and Harvey were noticeably frustrated. At this point, Harvey said, “It sounded good when he [the instructor] said it but...I think my brain quit,” and Earl added, “Holy crap, this is killing me.” When I realized that this group was still working on the first question, I asked what they had ruled out. Harvey replied that they had it down to A and B, and I explained the

mistake in A. The group erupted into a mixture of sighs, groans, and sardonic laughter as they realized that Tobias's initial explanation had been correct (in fact, Tobias's first utterance in several minutes of the recording was a prolonged groan). Unfortunately, neither listening to portions of their discussion nor answering Earl's question had helped me to understand their concerns with the question.

In both cases discussed here, I incorrectly assumed that I could understand the groups' interaction with each other from their interaction with me. Nate's apparent dismissal of my question caused me to assume that his group was not adequately addressing incorrect answers, while Earl's question led me to believe that his group had already finished discussing the first question. Although I was able to diagnose and resolve the concern for Earl's group during class, I did not discover my mistake with Nate until I was analyzing the data after the course had ended. During my analysis of the group discussions, I noted several other instances in which the students' ignorance of their own misunderstandings hindered the group process and impeded my ability to teach effectively.

Close analysis of group processes. The remainder of this chapter consists of a close analysis of one group's performance during a logic games discussion in lesson 7 of Phase II. For ease of reference, I have included the text of this game and a recommended solution in Figure 1.



Grouping/Linear Combination Games

Game #1: December 1994 Questions 12-17

An art teacher will schedule exactly six of eight lectures—fresco, history, lithography, naturalism, oils, pastels, sculpture, and watercolors—for three days—1, 2, and 3. There will be exactly two lectures each day—morning and afternoon. Scheduling is governed by the following conditions:

Day 2 is the only day for which oils can be scheduled.

Neither sculpture nor watercolors can be scheduled for the afternoon.

Neither oils nor pastels can be scheduled for the same day as lithography.

If pastels is scheduled for day 1 or day 2, then the lectures scheduled for the day immediately following pastels must be fresco and history, not necessarily in that order.

12. Which one of the following is an acceptable schedule of lectures for days 1, 2, and 3, respectively?

GBL List

(A) Morning: lithography, history, sculpture
Afternoon: pastels, fresco, naturalism

(B) Morning: naturalism, oils, fresco
Afternoon: lithography, pastels, history

(C) Morning: oils, history, naturalism
Afternoon: pastels, fresco, lithography

(D) Morning: sculpture, lithography, naturalism
Afternoon: watercolors, fresco, pastels

(E) Morning: sculpture, pastels, fresco
Afternoon: lithography, history, naturalism

$P_2 \rightarrow HF_3$

13. If lithography and fresco are scheduled for the afternoons of day 2 and day 3, respectively, which one of the following is a lecture that could be scheduled for the afternoon of day 1?

LCL

(A) history
(B) oils
(C) pastels
(D) sculpture
(E) watercolors

B, D, E = Not Laws

C = No P_1 because no HF (F is with L)

14. If lithography and history are scheduled for the mornings of day 2 and day 3, respectively, which one of the following lectures could be scheduled for the morning of day 1?

LCL

(A) fresco
(B) naturalism
(C) oils
(D) pastels
(E) sculpture

If L is 2, then O is out. Since you can only have two lectures out, then either S or W must be scheduled. That means S or W fills the last open morning spot.

P	N	F		N/F	F/N	P
S	L	H	or	S	L	H
1	2	3		1	2	3

15. If oils and lithography are scheduled for the mornings of day 2 and day 3, respectively, which one of the following CANNOT be scheduled for any day?

LCL

(A) fresco
(B) history
(C) naturalism
(D) pastels
(E) sculpture

If O and L are on days 2 and 3, then you can never have HF together on 2 or 3. That means you cannot have P on 1 or 2. And since P cannot go with L, P cannot be on 3 either. So P cannot be scheduled.

16. If neither fresco nor naturalism is scheduled for any day, which one of the following must be scheduled for day 1?

LCL

(A) history
(B) lithography
(C) oils
(D) pastels
(E) sculpture

If 2 lectures are out, the rest must be scheduled. That means O must be scheduled, and must be day 2. And since F is out, you cannot have HF on 2 or 3. So P must be scheduled for day 3. If O is 2 and P is 3, then L must be scheduled for day 1 (L cannot be with O or P).

17. If the lectures scheduled for the mornings are fresco, history, and lithography, not necessarily in that order, which one of the following could be true?

LCL

(A) Lithography is scheduled for day 3.
(B) Naturalism is scheduled for day 2.
(C) Fresco is scheduled for the same day as naturalism.
(D) History is scheduled for the same day as naturalism.
(E) History is scheduled for the same day as oils.

F H L in mornings means that S and W are out. So O must be day 2 PM. And since H and F are different days, then P cannot be 1 or 2. So P must be 3.

Schedule: N O P / L H/F F/H / 1 2 3

Lectures: F H L N O P S W[a]
Days: 1 2 3
Time: AM PM

PM ___ ___ ___ S W
AM ___ ___ ___
1 2 3
Ø Ø

O/L P/L

$P_{1,2} \rightarrow$ F/H H/F $_{2,3}$

Figure 1. Game 1 from Lesson 7.

This lesson occurred shortly after students had taken test 2. During this class, there were three groups, as shown in Table 10, which includes students' performance on the logic games section for both tests taken by that point:

Table 10. *Group Membership and Logic Games Performance on Tests 1 and 2 for Students in Lesson 7 of Phase II.*

	Test 1 LG	**Test 2 LG**	**Diff.**
Group A			
Danielle	48%	82%	+34%
Nate	43%	39%	-4%
Tobias	52%	39%	-13%
Group B			
Harvey	22%	39%	+17%
Ernesto	17%	4%	-13%
Don	35%	43%	+8
Group C			
Fiona	74%	69%	-5%
Debby	22%	39%	+17%
Earl	26%	56%	+30%

My adviser also attended and observed this lesson. The lesson focused on challenging logic games and the group members were asked to spend four minutes creating a setup for each game and then discuss the setup within their groups. Following the group discussion, I presented a possible setup for the whole class and then asked the students to spend six minutes solving the problems individually before discussing the answers in their groups.

Transcripts for all three groups were recorded and analyzed. Each conversation was analyzed for the number of substantive turns and the amount of backchanneling (i.e., non-substantive utterances or vocalizations, generally affirmations to keep the conversation flowing). Following Barron (2003), a turn was defined as a statement or set of continuous statements by one speaker. Crosstalk, during which multiple members

spoke simultaneously without interrupting any student's turn, was counted as a separate turn. If a turn was interrupted and the student stopped speaking, the turn was considered complete, whether or not the student returned to the same statement following the interruption. Table 11 shows the results of this analysis (note that I spoke to each group separately, as well to the entire class. Only those turns that were directed to a group were counted in this analysis):

Table 11. *Proportion of Substantive and Backchanneling Turns per Group Member in Lesson 7 of Phase II.*

	Substantive	**Backchanneling**
Group A	324	18
Danielle (% of total)	115 (35.5%)	1 (5.6%)
Nate (% of total)	65 (20%)	13 (72.2%)
Tobias (% of total)	75 (23.2%)	4 (22.2%)
Me (% of total)	69 (21.3%)	0 (0%)
Group B	133	29
Harvey (% of total)	63 (47.4%)	7 (24.1%)
Ernesto (% of total)	22 (16.5%)	18 (62.1%)
Don (% of total)	25 (18.8%)	3 (10.3%)
Me (% of total)	23 (17.3%)	1 (3.5%)
Group C	160	36
Fiona (% of total)	82 (51%)	4 (11%)
Debby (% of total)	40 (25%)	28 (78%)
Earl (% of total)	22 (14%)	4 (11%)
Me (% of total)	16 (10%)	0

Here, Danielle spoke 36% of the substantive turns during the conversation and only backchanneled during another students' turn once. By contrast, Nate spoke 20% of the substantive and backchanneled 13 times (72% of the total backchanneling within the group's discussion). During my analysis of these conversations, it was clear that each group had a dominant member and a conversational foil. In Group A, Danielle was the

dominant speaker and Nate was the primary foil.

Groups A and C were led by the student with the highest score from the second test (Danielle and Fiona, respectively). For Groups B and C, the dominant speaker contributed more substantive turns than both other group members combined, whereas Group A was slightly more balanced. During the recorded discussion, Groups A and B attempted five of the six questions, and Group C finished all six questions.

The turns indicated here for Group B are only during the six minutes of discussion about the questions (for the first four minutes of setup discussion, one of the group members had inadvertently shut off the tape recorder). Thus, it is likely that Group C had both the fewest turns and finished the most questions. This was perhaps in part due to the fact that Fiona had attempted this game previously and spent very little time talking about incorrect or inefficient approaches. Also, Earl was ill during this lesson and spoke far less than usual. Because Group B's discussion was incomplete and Group C's discussion consisted essentially of Fiona lecturing Earl and Debby, I have focused my analysis on Group A's discussion. Where applicable, excerpts from Group B and C have been included (in the form of extended block quotes) to provide contrast or confirmation of themes observed in Group A.

Discourse analysis of members of Group A. Relative to Group C, Danielle, Nate, and Tobias spent far more time discussing the setup and rules, with each group member offering significant contributions to the group's overall approach. As usual for this group, Danielle took the lead:

1. Danielle: Ok, so we know we start off with three days and two options. So would you guys…how would you guys pair it up?
2. Nate: So I did mine…
3. Danielle: Yeah, I did mine a linear version of that.
4. Nate: A linear version?
5. Danielle: It's the same thing. It's just what I'm more comfortable with.
6. Tobias: [*crosstalk*] What does yours look like? 7. Nate: [*crosstalk*] Can I see it? Are you hiding it? [*Danielle shows diagram to the group*] That's linear?
8. Danielle: Well, like a….
9. Tobias: [*interrupts*] I thought it was vertical.
10. Nate: [*crosstalk*] Oh, vertical, ok. 11. Danielle: [*crosstalk*] Vertical. That, not, sorry.
12. Nate: Confusion of terms here.
13. Danielle: Sorry. Instead of a horizontal, I did a vertical.

Unlike Fiona's group which simply followed along as she explained her diagram, Danielle, Nate and Tobias had a prolonged conversation about several different possible diagrams. The first convention they addressed was whether the diagram should be horizontal (two rows of three spaces each) or vertical (three rows of two spaces each). The typical convention for the class has been to use horizontal diagrams, but Nate and Tobias both knew that Danielle tended to score very highly in this section and wanted to know what she had done. In fact, in a previous lesson, Danielle had hidden her work from Nate so that he would talk about his own approach instead of simply mimicking her diagram. Here, Tobias's and Nate's eagerness to see Danielle's diagram likely resulted from recognition of her ability, the fact that she had hidden it from view, and her

comment that she "did a linear version of" Nate's diagram. However, after viewing her diagram, neither Nate nor Tobias changed from their horizontal diagrams and the conversation continued:

	14. Nate: So, uh, do you feel the need to show the out spaces? The spaces that are out, know what I mean?
15. Danielle: I feel like that would just confuse me, but it would be a good idea.	
	16. Nate: Well, I don't know. We did it earlier for the grouping, but this is like hybrid.
17. Danielle: It would be a good idea just to say that these two are out.	
18. Tobias: Because didn't he say if you know what's out, you automatically know what's in?	

The convention addressed above is whether or not to allot spaces for the unused variables. In the previous two lessons, I had repeatedly emphasized that a complete diagram should include spaces for all variables, used or not. In turn 15, Danielle offered a mixed endorsement of the strategy, saying both that it would confuse her and that "it would be a good idea." Nate mentioned that we had included an out group in similar games before, and Tobias paraphrased my justification for including it. This negotiation of meaning established early parity within the group as all three justified to themselves and each other the benefit of including an out group, and all three modified their diagrams afterward.

This is in stark contrast to Fiona's group, where Fiona led her group through the entire setup by the time I came around to speak with them. When I got to Group C, I noticed that all three group members were using a diagram that did not include an out group. I asked Fiona about this choice below:

20. Me: Do you have an out group?

[Fiona shrugs]

What I mean to say is you don't have an out group and you probably should.

21. Debby: Alright.

22. Me: A space to put the two variables.

23. Earl: These three are the in groups, you're saying? This is an in group and you gotta have…?

24. Debby: [*whispered*] I just added mine.

25. Fiona: I don't know.

26. Me: You don't want to have an out group?

27. Fiona: No.

28. Me: I'll show you why you should.

29. Fiona: I've done this game before, so….

As shown in Table 10, Fiona was among the best in the class at logic games (she had scored the highest of all students on the first test, and second highest on the second test). She had also prepared far more extensively than most other students prior to the course and, at times, openly disagreed with my suggestions. Although I was reasonably certain that Fiona could complete the game without changing her initial diagram, I was concerned that her intransigence would make the game harder for Earl and Debby:

30. Me: Ok, well, I'll explain why you should, even though…thanks, Fiona.

[*Fiona laughing*]

If F and N are not scheduled…

31. Debby: Oh ok.

30. Me: ...that's one example where you say I'm putting F and N and you should have a spot on your main diagram to say those two are out so you can put everything else in.

32. Fiona: Sorry.

33. Me: No, that's fine. In fact, you should always have an out group. You should never, ever do a game where there are more variables than spaces and not have room for all the variables. It's just good practice.

34. Earl: [*to me*] What's that?

35. Me: [*to Earl*] Always have an out group. If there's more variables than spaces. Your book doesn't and that's one of the things I actually don't like about it, is that it doesn't account for the out variables, but it should. You should make it balanced if you can.

36. Debby: [*to Fiona, quietly*] Whoa.

37. Fiona: Ok, we're done. [*stops recorder*]

Both Fiona and Debby seemed taken aback by my insistence. After reluctantly amending her diagram, Fiona abruptly stopped the recorder before Earl had completed the diagram. Leaving Earl a few steps behind was a very common theme for Group C.

Group A's discussion continued in turns 19-37 as Nate, Danielle, and I discussed exactly how to diagram the out group. However, Tobias had chosen a slightly different approach, which led to the following exchange:

38. Nate: Are we looking at the rules yet?
39. Tobias: [*crosstalk*] I thought it was just the diagram. 40. Danielle: [*crosstalk*] It's not two out in the morning and two out in the afternoon [*apparently referring to Tobias's diagram*]. There's just two out. There's only eight spots. Right now you have ten.
41. Tobias: Nah, I know I just put one at the bottom.
42. Danielle: No, that's going to be confusing.
43. Tobias: It's not going to be confusing.

Having established that an out group would be of value for all three group members, Nate asked if the group was ready to start discussing the rules together. Before they began, however, Danielle noticed that Tobias' diagram included two out spaces for the morning row and two more out spaces for the afternoon row. Danielle immediately pointed out this discrepancy in turn 40. The contrast between Danielle's nonchalance regarding differing orientations of the diagrams and her insistence that Tobias's out spaces are now

incorrect is striking and may have caught Tobias off-guard, as he spent the next several turns defensively maintaining that his diagram would not be confusing. It appears that Danielle's reaction here was borne of her belief that Tobias' alteration was a functional defect rather than a cosmetic difference, and her emphasis reflected concern for Tobias's ability to avoid being confused by his setup. Danielle pushed harder in turn 44, insisting that Tobias's diagram made it seem like there were ten spots:

44. Danielle: Because that, [*crosstalk*] that makes it look like there's ten. 45. Tobias: [*exasperated laughter, crosstalk*] I promise you, it's not. That makes it seem like there are ten spots. If you do one and one, it makes sense, but, like the morning doesn't have two spots out and the afternoon doesn't have two spots out. There's two spots out for the entire game. You see what I'm saying?
46. Tobias: [*defensively*] I see what you're saying.
47. Nate: You did that because you're trying to say, oh well, they could be in the morning or in the afternoon, the two extra people could be morning or afternoon….is that why you did that like that?
48. Tobias: Yeah, that's what I thought…I didn't really think that I need to add just two spots.
49. Nate: Well, that's what I was saying to Jason, but Jason was saying you only want to add two. It doesn't matter where you draw them. It doesn't matter morning or afternoon because they're just not there at all.
50. Danielle: [*crosstalk*] Yeah. 51. Tobias: [*crosstalk*] Just add two like that?
52. Danielle: There's just two out. We don't know if they're supposed to be out in the morning or out in the afternoon. [*crosstalk*] They're just out. 53. Nate: [*interrupts, crosstalk*] It doesn't matter, if they're out in the morning or the afternoon. [*crosstalk*] Exactly. If they're out, they're out. [*crosstalk*] What does matter is that you don't have four who are out. [*laughing*]

54. Danielle: Yeah. So, you have eight seats, six that are in, two out. There's no reason you want to put four spots there.
[*Tobias changes diagram*]

After Danielle's third attempt to convince Tobias failed, Nate used a different approach. Nate suggested a reason why Tobias may have chosen to draw his diagram with four spaces, "You did that because you're trying to say, oh well, they could be in the morning or in the afternoon, the two extra people could be morning or afternoon….is that why you did that like that?" and in turn 49, appealed to my authority rather than just arguing that Tobias was incorrect. Although Nate's eventual argument is identical to Danielle's, that morning and afternoon are irrelevant for determining out spaces, Tobias only began to consider amending his diagram after Nate's explanation. The exchange ended in turns 52-54 with Nate and Danielle skillfully reinforcing each other's efforts to help Tobias understand the correct approach and Tobias's decision to use just two spaces.

There are several possible explanations for the success of Nate's appeal after Danielle's failed attempts. Perhaps Tobias's resistance was diminished by Nate's sympathetic attempt to understand Tobias's approach or by the knowledge that both Nate and Danielle favored a different approach. It's also possible that Tobias listened because Nate mentioned me or because he perceived Nate to have more credibility than Danielle, although other observations from the classroom cause me to doubt either of these explanation. Regardless, these 53 turns represent the different degrees to which each group member was willing to change their approach upon realizing the others had approached the problem differently. When Jarret and Nate sought further input from Danielle about diagram orientation, no changes resulted to any group member's

approach. However, when Nate and Danielle offered unsolicited feedback to Tobias, he was eventually persuaded to adjust his approach. This exchange marks one of the clearest instances I saw of group work preventing confusion and yielding co-construction as the group members helped each other understand, justify, and properly implement an approach none of them were using prior to the discussion.

The group then turned to a discussion of the rules, jointly navigating the first two rules before diverging in their approach to rule three:

63. Tobias: [*to himself*] S and W in the afternoon…	64. Nate: Oohh, neither oils nor pastels can be scheduled for the same as lithography.
	65. Danielle: How would you go about that?
Neither oils nor…	66. Nate: I'll probably do like an L means not P, not O.
	67. Danielle: Or would you do an O or P means not L? Rather than doing L means not O or not P?
What the heck?	68. Me: Those mean the same thing, right?
	69. Danielle: It means the same thing?
O, P.	70. Me: If you have the presence of one means that both of the others are out.
	71. Danielle: Um-hm.
...day 1 or 2, then lectures scheduled for the day immediately following pastels…	72. Me: Then your contrapositive would be one or the other means L is out.
	73. Nate: Is that the bi-conditional?
	74. Me: Umm, it's a compound necessary condition.

What?! [*chuckling*]	75. Nate: Ok, compound.	
		76. Me: So if you said L means neither O nor P, and O or P means not L, those are identical.
	77. Nate: Ok.	
		78. Me: And both of them work.

The success of the previous interaction makes turns 63-78 especially striking, as Tobias, who had so clearly benefited from the group interaction, began working on his own instead of keeping pace with Nate and Danielle. In the first few seconds of turn 63, Tobias restated a portion of the second rule as Nate engaged the third rule. While Nate, Danielle and I had an evenly balanced discussion with several elaboration and clarification questions, Tobias read the rules aloud to himself and asked two rhetorical questions that are not picked up by any other group members. Although it is possible that Tobias was forced to work on his own because he fell behind the others' discussion, the fact that Tobias addressed three different rules during the time that Nate, Danielle, and I discussed a single rule suggests that Tobias was able to catch up and even surpass the others' pace. Thus, working alone seems to have been a deliberate strategy.

Although there were numerous other instances in this study of group members working in parallel rather than collaborating, this instance was instructive as it demonstrated that some students, even when confronted with evidence of the possible effectiveness of group work, would not fully commit to collaboration. Approximately 10% (8/79) of Tobias's turns were directed to himself and unconnected to the previous or succeeding turn of another group member (compared with 1% of Danielle's turns and 0% of Nate's turns). For most of the lesson, however, Tobias was engaged with his group,

only reading question stems and answer choices aloud to himself before joining Nate's and Danielle's ongoing conversation.

In contrast, in Group B, Earl was forced to work alone for nearly the duration of the game. During Group B's discussion, there were 66 conversational exchanges (defined as the set of turns and responses regarding a single topic, typically two to three turns long). Earl participated in 12 of the exchanges, two of them with me and 10 with Fiona. All 10 exchanges with Fiona and one of the two exchanges with me were initiated by Earl asking a question. At no point during the discussion did Earl and Debby address each other or present verbal evidence of any interaction (such as backchanneling each other's statements or taking up ideas presented by each other). Furthermore, the exchanges with Fiona consisted almost entirely of Earl asking her to explain something she had just discussed with Debby, as in the following:

> 141. Earl: What did you do for 16?
>
> 142. Fiona: Um, 16 was if F and N are both out, then P has to be on day 3, because otherwise the P rule won't work.
>
> 143. Earl: I got there's no O or no P, but…
>
> 144. Fiona: Yeah, and then we got lithography. No, there is O and P. On 16?
>
> 145. Earl: Yeah.
>
> 146. Fiona: Yeah, the ones that are out are F and N so everything else has to be in.
>
> 147. Earl: [rereading question stem to himself] "Which one of the following must be scheduled on day one"?
>
> 148. Fiona: Yeah, so P has to be on day 3 because otherwise its rule doesn't work.
>
> 149. Earl: Right.
>
> 150. Fiona: O has to be on day 2.
>
> 151. Earl: Right.
>
> 152. Fiona: L can't be on the same day as L or P, so it has to be on day 1.
>
> 153. Earl: So L's on 1?

154. Fiona: Yes.

155. Earl: So it's B?

156. Fiona: Yes.

Turn 144 here also contains the only instance during this game of Fiona asking Earl a question (in this case, a simple clarification question). Earl was never working on the same questions at the same time as Debby and Fiona, and despite my repeated efforts to encourage Fiona to work with him, Earl was only included in the dialogue when he asked questions of his own.

Returning to Group A, Tobias's choice to interpret the rules by himself rather than joining the conversation between Nate, Danielle, and me made it difficult for him to rejoin the conversation. In turns 79-84 below, Danielle and Nate tried to work out my instruction regarding the relationship between L, O, and P:

79. Danielle: But not… O or not P doesn't necessarily mean [*crosstalk*] L.	80. Nate: [*crosstalk*] L. Correct. So I don't know what he was saying.
81. Danielle: That's why I'm so confused.	
	82. Nate: Is that what he was saying?
83. Danielle: No.	
	84. Nate: No, that's not what he was saying.

Here, in turn 79, both Danielle and Nate appear to be confused about my statement (turn 76) that "if you said L means neither O nor P, and O or P means not L, those are identical." They correctly inferred that not scheduling O or P would not require L to be scheduled, but still could not understand what the rule meant. Turns 80-84 served only to confirm that neither student knew what I was saying.

In the following turn, Tobias, who had earlier asked himself "What the heck?" while working on this rule, tried to rejoin the conversation. However, Nate and Danielle were so preoccupied with understanding what I had said to them, that neither of them could answer his question. During the rest of the discussion, Tobias observed silently, either unwilling or unable to help Nate and Danielle understand the rule:

85. Tobias: So, how you doing that?	
	86. Danielle: I don't know. That's what I'm trying to figure out. You can't say…
	87. Nate: Isn't it just L means not O or P?
	88. Danielle: [*to herself*] So, if lithography is there…
	89. Nate: But O or P means not L. That's what he was saying.
	90. Danielle: O or P not L. Ok, that makes sense. O and P mean not L, the contrapositive.
	91. Nate: "O or" or "O and"?
	92. Danielle: Because the first one is not O or not P, right?
93. Tobias: [*to himself, inaudible*] So the contrapositive would be O and P means not L.	
	94. Nate: You sure?
	95. Danielle: No, I'm not, but that's the contrapositive. [*Danielle,* Nate *laughing*] Right, or am I doing something wrong?
	96. Nate: I mean, I know what you're trying to go by the way we did it before, like when you…
	97. Me: Ok, I'm going to go ahead and talk about the diagram that I would do.
	98. Danielle: No!!!
	99. Nate: Stop the recorder.

In turns 86-99, Danielle and Nate continued to discuss the relationship between L, O, and

P, but much to Danielle's apparent chagrin, were unable to resolve it before I ended the group work session. Based on Nate's and Danielle's responses to my explanation about this rule a few minutes earlier, I assumed that they both understood the implications and would be able to help Tobias catch up. As their discussion continued, however, it was clear that they struggled to use a consistent terminology for describing the relationship between the variables and lost track of the original wording of the rule. This was a common problem throughout the course, especially in the early lessons, as even proficient students were often lacking the confidence or ability to explain their insights or concerns to other group members. Even though Nate noticed that Danielle's summary in turn 90 differed from his inference in 89, Nate was uncertain whether she was wrong.

Additionally, Nate's question in turn 91 was precisely the right question to ask, but he did not understand what Danielle had done incorrectly in her answer. Turn 92 finally revealed that Danielle had never properly understood the initial rule. The rule states "Neither O nor P is scheduled for the same day as L." Danielle mistakenly believed "neither O nor P" to mean "not O or not P" instead of "not O and not P." The contrapositive of the correct interpretation is exactly as Nate stated in turn 89, but Danielle's understanding of the rule would have meant that one needed to schedule both O and P to eliminate L. Given that only two sessions could be scheduled per day, such a restriction would have been of little value, since it effectively stated that if a day is full with O and P, L cannot be on that same day.

Further complicating the interaction was Danielle's obvious skill in manipulating the logical conditions and her deserved reputation within the group. Inadvertently,

Danielle drew Nate's attention away from her use of an incorrect conjunction by incorrectly rephrasing the original rule and then correctly deriving the contrapositive of her rephrased statement. Throughout the course, I had emphasized repeatedly the importance of deriving contrapositives correctly, and Danielle was among the most competent at doing so in the class. Just as I occasionally overlooked errors by focusing on the wrong cues from my students, Nate (and probably Tobias) would have been attuned to the likelihood of errors arising in the contrapositive rather than in restating the rules. Once Danielle started to draw attention to the contrapositive, the group was no longer focused on the improper conjunctions.

Group B also struggled with this rule. Question 14 specifies that L and H are scheduled, and the members of Group B mistakenly interpreted this rule to mean that if L was scheduled, O and P could not be scheduled for any day, instead of not on the same day as L. Using this logic, Harvey and Don established that all of the remaining variables were included (besides O and P) and that S would have to be in the morning of day 1:

> 72. Harvey: So that means we know those two are out [O and P?], so we know the rest of them are in. Sculpture can't be in the afternoon, so S has to be in the morning, right?
>
> 73. Don: That's what I was leaning towards.
>
> 74. Harvey: [*to Ernesto*] Did you get that?
>
> 75. Ernesto: What's that?
>
> 76. Harvey: Well, look, L and H are in…
>
> 77. Ernesto: Yes, sir.
>
> 78. Harvey: So absolutely O and P are out, so we know those six are in, sculpture's in and it can't be in the afternoon, so it's good to go on day one in the morning.
>
> 79. Ernesto: Let's go with it.

80. Harvey: Let's go with it.

[*Ernesto laughing*]

However, the next question stem forced Harvey to reconsider his interpretation of this rule:

80. Harvey: Oils and lithography scheduled for mornings of day two and three…for the mornings of day two and three. Oils, lithography. Which of the following cannot be scheduled for any day?

P. If L, then not P or O. Wait…

81. Ernesto: Hmm?

82. Harvey: If L…

[*pause*]

Oh, for the same day as lithography…oh, see we might have messed that one up…

Here, Harvey realized that the question stem could not require an arrangement that violated the rules, and if O and L could be included for the same schedule, his rationale for the previous question was flawed. Unfortunately, before Harvey could fix his mistake, I interrupted him:

83. Me: [*interrupts*] You still working on 14?

84. Harvey: Well, I just realized we jumped to the…it was just the same day and not the entire series.

85. Me: Ok.

86. Harvey: So we might be wrong on that one.

87. Me: You got the right answer, so…

88. Harvey: Oh, good.

89. Me: Whatever your rationale was, you got the right answer.

90. Harvey: But this one, we can't use the same, because we were using the L rule, but it's just the same day.

Because I had missed their discussion about question 14, I assumed Harvey was having difficulty applying the rule to question 15 rather than rethinking how he had applied it to

question 14. Rather than pursuing Harvey's concern about question 14, I helped the group solve question 15, and it is unclear from the recording whether any of the group members later worked out the correct reasoning for question 14. Here again, the group work process revealed more to the students than it did to me.

Group A was the only group that had continued discussing the setup until I ended that group work segment. During the six minute discussion of questions, Group A solved questions 12 and 13 with little difficulty, prompting the following exchange:

	156. Nate: This isn't that bad.
157. Danielle: If you understand the rules, yeah.	
	158. Nate: If Danielle wasn't here, I might not be doing so well.
159. Danielle: Ok, then I'm going to stop talking. You lead the discussion.	
	160. Nate: Oh, yeah. [*laughing*] Alright, Miss Teacher.
161. Danielle: Go for it.	

Though lighthearted, this conversation revealed Danielle's genuine concern that Nate was too reliant on her for solving games (conversely, in logical reasoning, Danielle was far more likely to seek Nate's help than vice versa). Gamely, Nate tried to lead the discussion of question 14, while Tobias unsurprisingly began the question on his own. Also unsurprisingly, Danielle was unable to cede the floor to Nate for long, ignoring Nate's suggestion to start drawing and reading the rest of the question aloud to herself. This led to the amusing circumstance of all three students reading the same question aloud within ten seconds of each other.

Once Danielle had finished reading the question, the group was able to eliminate

two of the five answers before getting stuck. Rather than solving the question for them, I explained what the group should do if they faced this situation on test day:

	193. Me: …so at this point if you don't see any other rules that are being obviously broken, then it's time to test. So put F in the morning of day one and see if you can create a diagram that works that way. You've got your morning filled in, think about the five that are remaining.
194. Tobias: We've got…	
	195. Me: O P S W.
196. Tobias: Can't have S or W.	
	197. Me: Yeah, so did you catch that? You can't have S and W at all now.

The group had previously eliminated O and P, and was trying to decide between F, N, and S. Tobias quickly determined that including F in the morning would force out too many variables to have a complete schedule, but neither he nor the other group members grasped that this meant a valid setup would have to include either S or W.

In the following turns, I finished explaining why N could not be correct either and then readdressed the problem from the beginning:

	212. Me: Yeah. And I'll explain why. With L on day two that means O is out. Right?
213. Tobias: Um-hm.	
	214. Me: If you start with that, that O is definitely out…
215. Nate: [*crosstalk*] Um-hm.	

	[*crosstalk*] …that will funnel you toward the correct answer more quickly. So, L in on day two, means O out for sure. Does that make sense? If L's on day two, then O's out? So that means that you only have one more space out, total. And you look at the things that are left, that you could include…

After a digression to encourage Nate to draw an out group, I suggested that the group should be able to see the constraint now and added, "And if you can't, Danielle will actually explain it now, even if she doesn't want to."

At this point, I mistakenly believed that Danielle understood the rules sufficiently to explain the solution to this problem. Unfortunately, Danielle was still confused regarding the relationship between L, O, and P, which caused some confusion during her explanation:

	241. Danielle: So if you have O in, in A and B, that means you'd have to have three out, but we only have two out spots. C is the only one, where since S is already in, it's not taking up one of those out spots.
	242. Nate: Um-hm.
	243. Danielle: So O and W are the out spots.
244. Tobias: Good lord.	
	245. Danielle: Does that actually make sense or are you just saying yes?
	246. Nate: No, I see. I see why…so you said S L H F N P?
	247. Danielle: Um-hm. [*crosstalk*] N P F.
248: Tobias: [*crosstalk*] N P F. Because P has to be on day 1 or 2, so the third day has to be H and F.	
	249. Nate: Oh, ok, yeah that makes sense.

Danielle was correct that O and W were the out spots, but her hypothetical example was flawed because L precludes O from being included at all, regardless of how many variables would otherwise be out. Despite Danielle's confident tone and my endorsement of her understanding, Nate and Tobias were justifiably confused. In 246, Nate either feigned understanding or misheard the explanation because Danielle's actual explanation is incomprehensible. In restating the potential solution, Nate also fixed a mistake that both Danielle and Tobias had overlooked by moving P off of day 2 into day 3, where it would no longer conflict with L. Unfortunately, both Tobias and Danielle "fixed" this correction by reintroducing the earlier error, followed by another incorrect explanation from Danielle.

Like her class mates, I assumed Danielle understood more than she did. She had led the group to answer the previous questions correctly and appeared to understand my explanation to the group. One of the consistently frustrating and surprising results of my analysis here was discovering how often I mistook confidence for comprehension, especially when a student had chosen the correct answer for the wrong reasons. The problem was particularly prevalent during this game, as all three groups used faulty logic to get the correct answer to at least one problem. And each time, I was so focused on the outcome rather than the process that I did not notice the error until listening to the transcripts.

Although Danielle struggled with problem 14, she solved question 15 almost instantaneously:

251. Danielle: [*crosstalk, to group*] Ok, 15. If oils and lithography are scheduled for the mornings of day 2 and 3, respectively, which one of the following CANNOT be scheduled for any day? Just off the bat, I would think P if I was low because O is already on day 2 which means F and H can't go on day 2 or 3.
252. Tobias: P can't be in the morning…
253. Danielle: Right? So P can't go on day one because F and H can't go anywhere else, so the only [place] where P could go would be underneath L, but L and P can't be on the same day. That was just off the bat for me. Does that make sense how I looked at that? I don't, I don't even know if it's right, but that's how I looked at it.
254. Nate: P can't be on day 2?
255. Danielle: If P is on day 2, then F and H have to be on day 3, but L is already taking up one of the third spots.
256. Nate: Oh, yeah.
257. Danielle: Make sense? 258. Tobias: [*crosstalk, to himself*] P can't be after L. [*crosstalk, to Nate*] So P can't be on day 1 or 2. 259. Tobias: [*interrupts, crosstalk, to group*] P can't be on [inaudible], either. [*crosstalk, to group*] And if P goes on the third day…if P goes on the third day, then it's the L and P rule that it's breaking. [*to me*] That was really creepy. I literally looked at it and I was like, P.

On this problem, Danielle's approach was instinctive and accurate. Her explanation was also clear and precise. Most interestingly and confusingly, in turn 257, she correctly described the L and P rule that she had improperly construed on question 14.

Altogether, Group A attempted five of the six questions together and correctly solved four of them. The group members consistently showed respect for each other, responded to each other's questions, and maintained greater conversational balance than either of the other groups. What is perhaps most interesting about this conversation is the extent to which it reflected several of the themes found across the discussion from both

phases. Theme 2, reinforcing new concepts, and theme 3, group members identifying and correcting their own mistakes, were both present in this discussion, particularly as Danielle and Nate attempted to fix Tobias's initial diagram. This group also showed an unusual variation of theme 4, lack of cohesion reducing group effectiveness. Although Tobias sometimes deliberately worked alone and reduced the group's overall cohesion, his actions did not disrupt group effectiveness, and he offered several significant contributions to the group's problem-solving process. I also observed theme 6, mismatches between student-instructor interaction and student-student interaction, as I incorrectly assumed that Danielle would be able to relay my explanation to her groupmates. Thus, the discourse analysis reveals a fractal nature to the overall course experience, as many of the macro themes seen across group and courses were repeated on micro scale within a single group discussing a single game.

Summary of Results

Through analysis of both quantitative and qualitative data, I was able to develop a rich picture of the students' experience learning through a flipped curriculum. However, the implications of these results are complex and in some ways, ambiguous. These implications, along with qualified answers to the research questions of this study, will be discussed further in the following chapter.

Chapter Five: Discussion

Before I address my results and discuss the answers to the research questions, I will discuss the limitations I faced in conducting this study. I then will address my findings, and implications for practice and research.

Limitations

As with any study, this study had limitations that warrant caution with respect to generalizing findings and to claiming strong intervention effects. Many of the limitations are due simply to the nature of classroom research. For instance, many more students consented to participate in the research than actually completed the instruments given. Thirty-eight students completed the initial exam, of whom thirty-four (89%) consented to participate in the study. Only 11 students (32%) completed all three surveys and finished all four exams, and just two (6%) of these students attended all 13 class sessions. Thus, all of the analyses had lower than expected power. Although I had expected to have between 45 and 60 students across the three classes and had anticipated participation rates closer to 60%, concerns about power and sample size were a major consideration in adopting a design research focus.

Although design research is focused primarily on how theoretical frameworks can be iteratively deployed within a single learning environment, I had nonetheless hoped to gather baseline comparison data for measures of student learning and student engagement by obtaining data from students in other courses. These efforts failed and thus I could not administer the Student Course Engagement Questionnaire to students in a traditional course. As an alternative, I resorted to using my own previous courses to act as a

comparison for overall score results and attendance, as these courses served students from a similar population and were taught by the same instructor in the same location with the same course materials. Thus, the lack of a true control group for some of the research measures, although necessitating revision of some of the original research questions, does not prevent this study from offering meaningful results about some of the ways in which traditional and flipped courses differ from each other.

Other limitations arise from the nature of design research itself. Among these limitations is the inability, in most cases, to isolate and control for a single variable and to assign students randomly to experimental conditions (Collins, Joseph, & Bielaczyc, 2004). As a result, design research generally involves gathering extensive amounts of quantitative and qualitative data. In this study, I captured more than 350 hours of audio recordings from both phases, in addition to field observations, quiz reports for 39 sessions, test data, survey responses, and semi-structured interviews. It quickly became apparent that I would be unable to analyze all of the data within the timeframe of this study. Thus, I decided to focus my quantitative analysis on six sets of recordings- lessons 2, 5, and 6 from Phase I, and lessons 2, 5, and 7 from Phase II. Each lesson contained recordings from three to four groups, which were transcribed, coded, and analyzed by me alone, with a colleague reviewing portions of the transcripts to help confirm themes. These recordings were chosen because the lessons on those days contained material that was commonly misunderstood by students and that I believed would result in interesting discussions (I had initially intended to include lesson 7 from both phases, but some of the group recordings from this lesson in Phase I were lost). Given more time or more

assistance from other researchers, it may have been possible to analyze more of the data, but I am convinced that the themes that arose and that I presented in Chapter Four are representative of those that would have been found elsewhere in the data.

Discussing Findings

As is befitting an investigation of a broad classroom intervention, the findings were complex and in some ways, equivocal. On the one hand, flipping the LSAT course helped me learn a great deal about how my students learn and how I might improve my teaching. It also helped create a more social and comfortable classroom environment. On the other hand, the data showed little or no effect on learning, and the environment may have been too casual in certain aspects. It also required a tremendous, one-time investment of time and effort to create the video lectures and adjust the syllabus, as well as continual adjustment to coaching individuals instead of lecturing the entire class. The following sections discuss these findings in greater detail organized by each of the research questions posed in Chapter One.

Effect of flipped curriculum on learning, course persistence, and engagement. The quantitative results regarding effect of flipped curriculum on student learning were mixed. A t-test of gain scores between the flipped courses and traditional courses revealed no significant differences on overall score or section subscores. However, Hake (1998) has suggested the use of normalized gain scores to measure pre/post-test changes for measuring individual change due to an intervention, where normalized scores are the actual gain scores expressed as a proportion of maximum possible gain. For this study, the maximum possible gain for overall (subsection) score is

the difference between 180 (100%) and the student's initial score (percent correct), and the actual gain score is the difference between the highest score (highest percent correct) and the student's initial score (percent correct).

Using normalized scores shows that students in traditional and flipped instruction differed significantly in overall gain ($t(57) = 1.66$, $p < .05$), with students in flipped courses ($M = .28$, $SD = .16$) realizing a higher proportion of their maximum possible score increase than students in traditional courses ($M = .20$, $SD = .18$). Subsection analysis indicated that students in traditional and flipped instruction did not significantly differ in normalized gain scores for either logical reasoning or reading comprehension, but did differ significantly in logic games ($t(57) = 1.71$, $p < .05$), with students in flipped courses achieving, on average, 40% of the maximum possible gain ($SD = .32$), compared to 28% ($SD = .36$) for students in traditional courses. These results suggest that a flipped course may have helped some students fulfill more of their potential on the exam and on logic games, in particular, compared to a traditional course. Because the actual gain scores ($M = 8$ for students in flipped courses and $M = 6.2$ for students in traditional courses) did not differ, the normalized gain scores may suggest an interaction between initial score and score increase. These data were also analyzed with ANCOVA, using pre-test scores as the covariate, and the results did not differ substantially.

Regarding attendance and engagement, I had expected the improved classroom atmosphere associated with a flipped course to be reflected in increased attendance and student engagement. However, attendance in Phase I was significantly lower than that in either traditional courses or Phase II, neither of which differed from each other. At best,

then, inverting the curriculum and introducing small group work to these courses had no effect on attendance for the set of flipped courses (although Phase I was the worst-attended course of the nine courses for which I have taken attendance during the last seven years).

Another unanticipated feature of attendance in the flipped courses was the increased prevalence of students choosing to leave class early. Although this behavior had certainly been present in traditional courses, it was unusual for even two students to leave during the same class. In contrast, four Phase I students left class during the break for lesson 11, and three students left class during two other lessons. Two students also left lessons during Phase II with some frequency. It is possible that some students were so averse to the flipped curriculum process that they sought to limit the amount of time they spent in that environment. However, students who left lessons early nearly always chose to leave before reading comprehension instruction, and this section of the course was presented nearly identically in both flipped and traditional courses (i.e., there was no group work during the reading comprehension section and very little video time was devoted to reading comprehension lectures). Conversely, my approach to reading comprehension may have given these students the impression that reading comprehension instruction was either of less importance or lower quality than the sections that were flipped during this study. In either case, this pattern was one of the more surprising observations of this study.

Experience for the instructor and students. Teaching a flipped class was revealing in many ways. For example, I was pleasantly surprised to find that nearly all of

the students who came to a session had prepared by either watching the video, reading the lesson text, or both. It was also rewarding to note that the increased degree of student-student and teacher-student interaction led to a more productive classroom atmosphere. I noticed that before class and during breaks, students in these courses were considerably more talkative and social with each other than previous courses had been. Students also seemed to be more accepting of conversational tangents and showed few, if any, signs of impatience or annoyance with me or with each other when conversation briefly strayed from LSAT topics. This is in marked contrast to traditional courses, during which off-topic comments by myself or students were frequently met with non-verbal indicators of frustration or impatience (and, on occasion, with comments such as, "Can we please get on with it?" or "What does this have to do with the LSAT?"). Furthermore, several friendships were formed during these courses, some of which endured well beyond the end of the course.

At times, I found the process of rotating from group to group to be a very effective way to diagnose student comprehension. For example, after listening to only a few turns of a conversation, I could often determine what students understood correctly and what they had overlooked. These impressions were later confirmed during my analysis of the group recordings. More significant were the occasions where the group discussion increased my understanding of the content. I often observed or overheard a student's approach to a problem that was more efficient or less complicated than my own. In a traditional course, such insights were only yielded when a student realized their approach differed from mine and was willing to ask about the differences. In a classroom

using a flipped curriculum, particularly during the logic games instruction, I could see each student's approach, and I routinely complimented students on their approaches, sometimes while addressing the whole class.

Using a flipped curriculum also helped me better understand my own strengths and weaknesses as an instructor. In my courses, I always attempt to answer students' questions as thoroughly and accurately as I can. Listening to the group discussions revealed that my answers were typically clear enough for students to take up the ideas in their own discussions after I had left their groups. However, there were numerous instances in the discussions when students assured me that my answer made sense and moments later, expressed confusion to one another. This suggested that rather than simply asking students whether or not an idea made sense, I should ask my students to explain the concept back to me in their own words. I also observed that my initial question to most groups (generally some variations of, "How's it going?" or "What are you working on?") was inadequate for revealing the groups' understanding of a problem. My interactions with the groups were generally more productive when I listened to the conversation for several seconds before joining in or when I asked more specific questions, such as, "How were you able to eliminate answer choice B?".

Perhaps the most difficult classroom management challenge presented by flipping the classroom was the pace of instruction. Students in high-achieving, complacent groups were often done discussing problems well before I could meet with each of the groups. These students often assumed that answering the questions correctly obviated the need for discussing the process. During Phase II, I encouraged students in groups that had

answered all of the questions correctly to appoint a different spokesperson per question who was responsible for explaining their approach to the group. However, the response to this suggestion was haphazard and uneven, with some groups ignoring my recommendation altogether and others allowing the same person to explain each of the questions.

The issue of properly pacing all groups was exacerbated when I attempted to resolve a challenging misconception during group discussions. At times, it was clear that members of groups that had finished quickly were listening to my explanation to another group. When I noticed this, I would briefly reset the issue and speak more loudly, which created issues of its own when two or more groups were listening to my explanation while the remaining groups continued their own discussion. Generally speaking, even after both phases and all three courses, I never felt satisfied that I had adequately dealt with the different paces of students among and within groups.

Whereas properly pacing group work was my primary frustration with the flipped curriculum, the survey responses, field observations, and interview data led me to believe that the students' main frustration was the opportunity cost of group work at the perceived expense of learning from my expertise. These issues are not altogether unrelated, as the students' sense that they were not learning enough from me was likely to be particularly acute when they were "finished" with group work before I had reconvened collective instruction. Complacency, frustration, overconfidence, and lack of commitment to recommended group processes could each result in groups prematurely ending discussion and passively waiting for me to start "teaching" again. Without recourse to

grades or some other method of creating group-level accountability, this perception of opportunity cost may be an intractable problem for implementing a flipped curriculum design for educational settings similar to test preparation classrooms.

Findings from group discourse. Group discourse revealed that the flipped curriculum could have a significantly differential impact on different groups of students. For students of various abilities, complacency and over-confidence could severely impede group function. The most effective group discussions included high-scoring students who were willing to question their own assumptions and were comfortable discussing both process and result. Many groups featured unbalanced discourse, with one student dominating the discussion, one student serving primarily as a foil for the dominant voice, and the remaining students participating in very few conversational turns. Among groups of four students, there were occasionally two dyads of students working in parallel with no dominant student, but it was unusual for the most vocal student in any group to have fewer than twice as many utterances as the least vocal, and ratios of 3:1 or greater were not uncommon. The imbalance would have been even more severe if measured in terms of time spent speaking or number of words uttered, rather than utterances.

Barron (2003) observed that the management of joint attention was crucial to effective group function and this study supports her claim. When even one member of a group was not paying attention to the others' conversation, it could result in cumbersome repetition, broken turns, unattended questions, and aborted explanations. In some cases, mismanagement of joint attention even undermined successful prior work the other group

members had accomplished. Although some students, such as Tobias, could alternate between working alone and working jointly without disrupting the overall group process, most students were more like Alberto or Colin, whose untimely questions delayed or derailed otherwise effective group discussion.

Referencing Schwartz (1995) and Shirouzu, Miyake, and Masukawa (2002), Barron (2003) further wrote, "There is even some intriguing experimental evidence that collaborators can generate strategies and abstract problem representations that are extremely unlikely to be observed when individuals work alone" (p. 309). In the six lessons I closely analyzed (three each from Phase I and Phase II), I noted several instances of students explaining a concept vaguely or hesitantly at first, and then with increasing articulation and clarity as they continued speaking. In these cases, the presence of listeners, whether or not those listeners contributed to the content of the explanation, afforded a problem-solving space within which knowledge arose. On one occasion, I could hear Kyle explain something incorrectly, react to group members' confusion, correct his own error, and explain his new understanding so effectively that the other group members were able to understand and apply it immediately. This was a striking, though hopefully not unique, instance of co-construction of knowledge that demonstrated the potential of collaborative work in educational settings similar to this one.

Implications for Practice

Within the context of design research, it is appropriate to discuss both suggestions for other practitioners and possible additional iterations of the current design. Thus, this section will address what characteristics would be present if I had gone on to a Phase III

of my instruction. I will then offer guidance to other instructors who may wish to implement a flipped curriculum with their students.

The next phase of this study would differ from the two flipped instruction phases reported here in several ways. First, I would use a more purposeful group composition process. In this study, the issue of group composition was essentially ceded to the students in phase I. Group sizes of either three or four were typical, and although there were several relatively stable dyads in the course, the specific groups varied from session to session due to attendance, seating, and shifting social dynamics (e.g., emerging friendships or annoyance with unresponsive group members). My response was to eliminate student choice altogether for Phase II and randomly assign students to groups for each class meeting. Although this second approach did yield some benefits in terms of classroom atmosphere and had no negative effect on scores, engagement, or perceived team effectiveness, a future iteration of this curriculum might benefit from a more purposeful group composition approach.

I considered homogeneous groups based on pre-test scores, but rejected this approach due to concerns about stigmatization and diminished motivation for lower scoring groups. Furthermore, research suggests that heterogeneous groups benefit low scorers to a greater degree than they are detrimental to high scorers (Webb, Nemer, Chizhik, & Sugrue, 1998), and heterogeneous grouping would likely result in increased opportunities for peer tutoring. Groups could be formed on the basis of test 1 scores, with the mean scores for each group selected to be roughly equivalent. A secondary consideration would be to distribute students with high prior knowledge (i.e., those who

had taken another prep course or had extensively reviewed prep books prior to the course) as evenly as possible among the groups.

To improve the likelihood of gathering meaningful data at the group level, the students should either remain in stable groups throughout the course or, possibly, be reassigned to new groups following test two. Michaelsen and Sweet (2008) identified three obstacles to successful group function: a) unequal distribution of member resources among groups, b) coalitions within groups, and c) group instability. My decision to allow students to select their own groups in Phase I was motivated by the belief that students would be more committed to group work if they had some choice in their group members. However, these groups fell victim to both unequal distribution of member resources and detrimental coalitions within groups, as pre-existing and emerging dyads of students with similar ability levels impeded overall group effectiveness and cohesion.

Randomly varying the groups in Phase II largely alleviated issues with group coalitions and also allowed students to partner with most of their peers for at least one lesson during the course. Although I believe this contributed to the generally improved classroom atmosphere, it was also a challenge for some students to "get used" to different group members and dynamics for each lesson. Random group composition also resulted, at least on occasion, in homogenized groups, which made pacing the lessons difficult. On balance, I am not convinced that the benefits to the classroom environment outweigh the expected gains of stable groups, especially given that even the more stable, self-selected groups in Phase I resulted in a better social environment than many of the traditional classrooms I had previously taught.

Another benefit of more stable groups is that the Team Beliefs and Behaviors Questionnaire and similar instruments are easier for students to understand and researchers to interpret when students are evaluating a single group rather than several different groups at once. Whether groups remain stable for the entire course or are rearranged at an appropriate intermediate point, I recommend administering team evaluation instruments mid-course and post-course, and utilizing the mid-course results as formative evaluation. Instructors may use these results to offer guidance to some students on more effective participation within permanent groups or to help make informed decisions about group balance and resource distribution for rearranging groups. Finally, repeated administrations of the team evaluation instruments would allow analysis of group function over time or comparison of different groups within the same course environment.

I also agree with Strayer's (2012) assertion that the out of classroom content must be well integrated with the classroom content for a flipped curriculum to be effective. In this study, the video lectures were entirely consistent with the classroom content but less coordinated than they might have been. In Phase I, due in part to force of habit and also to concerns that students might have come to class unprepared, I tended to deliver mini-lectures at the beginning of each section that often repeated portions of the video lectures. As I learned that students were, in general, both reading the lesson text and watching the videos, and that they often did so immediately before class (in fact, it was quite common for me to see students in the hallways or lobby outside of the classroom watching a video when I arrived for class), I reduced my lecture to address only those issues that arose

from the readiness assurance quizzes. Other practitioners would be well-advised to determine if their students are similarly compliant with the pre-class preparation process, and adjust the amount of in-class lecture accordingly.

Also, I would strongly encourage other instructors to plan how they may deal with differences in group pace. Carefully constructed, heterogeneous groups may reduce some of these differences, as may open-ended, ill-structured problems. In my courses, I would be more insistent that when students in a group all select the same answer, one person be designated to explain his or her approach. I would explain the value of articulating one's own thought process, emphasize the possibility that hearing one student's approach could improve or refine how the remaining students address the problem, and encourage students in these groups to rotate who is speaking for each unanimous question.

When, as in these courses, the problems used in a flipped curriculum have a definitively correct answer, the instructor should consider how and when to reveal what the correct answer is to students. In this study, students who learned which answer was correct too early often stopped discussing the problem without fully considering the incorrect answers or evaluating the effectiveness of various approaches. On the other hand, when I withheld the correct answer from groups until the discussions were over, some students spent significant time and effort in defending wrong answers to each other, resulting in obvious frustration. One approach I deemed satisfactory was to ask students to form an initial consensus before using the Socrative classroom response system to check their answer. To reduce the likelihood of students using personal electronic devices for diversionary activities, I asked for only one response per group, believing that the

person whose task it was to submit answers for the entire group would feel an obligation to keep up with the discussion and that the other students would have less reason to use their devices at all. Also, the consensus-seeking process allowed me to overhear discussion, and the group answers helped me determine which students needed the most instruction. Here again it is important not to let students simply move on to the next problem if their consensus answer is the correct answer.

Finally, I would put additional emphasis on outlining the expected benefits of collaboration in order to increase student commitment. Although I mentioned to my students what I hoped would happen during their collaboration with each other, I did not give them any evidence of the potential value of group work nor could I provide examples of effective and ineffective group function prior to this study. Providing a more explicit rationale for introducing group work into courses such as this one and offering practical direction for helping students make their group work more effective are recommended.

Implications for Further Research

Although this research does not unequivocally support the use of flipped curriculum designs for learners in test preparation courses or similar educational settings, it does warrant further investigation. The most apparent opportunity for future research to expand upon the results of this study is to provide more comparable control data for measures of learning and student engagement. While random assignment of students to flipped courses and traditional courses is not consistent with the ethos and constraints of design research, observing, measuring, and richly describing the learning atmosphere in a

traditional course would add to these findings here.

An intriguing possibility that the normalized gain score analysis suggests is that the flipped curriculum was more beneficial to students with higher initial scores (i.e., those for whom a given score increase represent a higher proportion of the total possible increase) than to students with lower initial scores. Due to sample size limitations here, I was unable to adequately investigate this possibility. Further research with a greater sample may then reveal whether an interaction exists between course design and initial test score, even in the absence of conclusive evidence from this study for a main effect of course design on student learning.

Another area of potential future research would be related to student expectations and beliefs about the learning environment. For example, research on persistence in adult education suggests that the gap between expectations and reality in the classroom may contribute significantly to reduced likelihood of course completion (Dirkx & Jha, 1994; Perin & Greenberg, 1994). Understanding what students expect—from the course, the instructors, themselves, and each other—may help explain differences in commitment to and belief in the group work processes investigated here. Finally, the role of incentives or rewards as a means of increasing group accountability may also reveal whether an intervention of this kind can be implemented more successfully among self-focused learners.

All of these areas are worthy of further investigation and may help answer questions raised by this study about the ideal roles of instructors and students, the appropriate use of time inside the classroom and outside of the classroom, and the myriad

ways in which group work can facilitate and impede learning of complex content among adult learners. Whether the benefits of a flipped curriculum justify its costs depends to a large degree on the answers to these questions and on an instructor's belief in their importance.

Appendix A

LSAT Course Syllabus

Lesson One

Logical Reasoning

Attacking the Stimulus
Premises, Conclusions, and Fact Sets
The Importance of Language
Analyzing the Question Stem
Logical Reasoning Question Types
Prephrasing Answers
Must Be True Questions
The Fact Test™
Main Point Questions

Logic Games

General Introduction
Setups and Diagramming
Pure Sequencing Diagramming Guidelines
Pure Sequencing Games

Lesson Two

Logic Games/Logical Reasoning

Logical Opposition and Certainty
Logical Reasoning
Must Be True Question Review
Sufficient and Necessary Conditions
Reversals and Negations
Chain Relationships
Diagramming "Unless"
Multiple Sufficient and Necessary Conditions
The Double Arrow

Reading Comprehension

The Two Passage Types
Approaching the Passages--7 Critical Steps
Using VIEWSTAMP™
Reading Comprehension Question Types
Attacking the Questions

Lesson Three

Logical Reasoning

Weaken Questions
Typical Weaken Scenarios
Three Incorrect Answer Traps
Cause and Effect Reasoning
How to Attack a Causal Conclusion
Causality in Weaken Questions

Logic Games

Basic Linear Setups
The Importance of Numbers
Rule Representation: Blocks, Sequencing, and Dual Options
Basic Linear Setup Games

Lesson Four

Logic Games

Advanced Linear Setups and Multiple Stacks
Repeated Variable Sets
Diagramming with Multiple Stacks

Logical Reasoning

Strengthen Questions
Causality and Strengthen Questions
Justify the Conclusion™ Questions
The Justify Formula™
Solving Justify Questions Mechanistically

Lesson Five

Logical Reasoning

Assumption Questions
The Supporter/Defender Assumption Model™
The Assumption Negation Technique™

Logic Games

The Principle of Grouping
Unified Grouping Theory™
The Double-Not Arrow
Linear vs Grouping Symbolizations
Hurdle the Uncertainty™
Defined Grouping Games

Lesson Six

Logic Games

Undefined and Partially Defined Grouping Games

Reading Comprehension

Diagramming the Passages
Passage Notations
Diversity Passages
Three Types of Diversity Passages

Lesson Seven

Logical Reasoning

Method of Reasoning Questions
Method-AP Questions
Fallacious Methods of Reasoning Categorized
Flaw in the Reasoning Questions

Logic Games

Grouping/Linear Combination Games
Working with the Combination of Major Principles

Lesson Eight

Logical Reasoning

Parallel Reasoning Questions
The Parallel Reasoning Elemental Attack™
Parallel Flaw Questions

Reading Comprehension

Comparative Reading Theory
Similarities and Differences
Comparative Reading Passages

Lesson Nine

Logical Reasoning

Numbers and Percentages--Common Misconceptions
Numbers and Percentages Questions

Logic Games

Numerical Distribution Games
Fixed versus Unfixed Distributions
Distribution Identification and Production Methodology™
Limited Solution Set Games
Identify the Templates™
The Dangers of Misapplication

Lesson Ten

Logical Reasoning

Principle Questions
Point at Issue Questions
Incorrect Answers in Point at Issue Questions
The Agree/Disagree Test™
Point of Agreement Questions
The Agree/Agree Test™

Reading Comprehension

Science and Technology Passages
Types of Science Passages
Handling Scientific Elements

Lesson Eleven

Logical Reasoning

Resolve the Paradox Questions
Evaluate the Argument Questions
The Variance Test™
Cannot Be True Questions

Logic Games

Logic Games Review
Killer Games

Lesson Twelve

Logical Reasoning

Logical Reasoning Review
Advanced Sufficient and Necessary Review Problems
Cause and Effect Review Problems

Reading Comprehension

Reading Comprehension Review
Law Related Passages
Two Special Topics

Appendix B

Pre-course Survey (control and experimental students)

What are your full initials?

Who is your instructor? Please use instructor's full name.

What is your goal score?

What schools do you plan to apply to? (Please list all schools that you are seriously considering)

What is your overall scaled score for practice test 1? Please enter total score only.

Using the results from the PowerScore online student center, please the following results for practice test 1 below. Refer to graphics for specific results needed.

	Number correct	Number incorrect
Logical Reasoning (must total 51)		
Sufficient and Necessary problems (must total 6)		
Reading Comprehension (must total 27)		
Logic Games (must total 23)		

What is your age?

What is your gender?

- Male
- Female

What is your undergraduate major?

What is your undergraduate GPA?

- Below 2.0
- Above 2.0 - 2.5
- Above 2.5 - 3.0
- Above 3.0 - 3.5
- Above 3.5 - 4.0
- Above 4.0

Who is paying for this course? This question is optional

- Self
- Parents
- Both
- Other _____

Please include any questions or comments you have about this survey below. Thank for your participation.

Appendix C

Mid-course survey (experimental students only)

This questionnaire examines your attitude to teamwork. Please indicate to what extent you agree with the following statements concerning the team in which you are working and the task with which you are dealing. Each statement has a 5-point rating scale; 1 = I disagree completely; 3 = neutral and 5 = I agree completely.

What are your full initials?

Items measuring interdependence

	I disagree completely	I disagree somewhat	Neutral	I agree somewhat	I agree completely
My group members depend on me for information and advice.					
I depend on my group members for information and advice.					
The group members agree on what we want to accomplish.					
When my group members succeed in their jobs, it works out positively for me.					

Items measuring social cohesion

	I disagree completely	I disagree somewhat	Neutral	I agree somewhat	I agree completely
I like my class.					
I get along with members of my class.					

I feel a sense of belongingness to my class.					
I am friends with the members of my class.					

Items measuring task cohesion

	I disagree completely	I disagree somewhat	Neutral	I agree somewhat	I agree completely
This team is united in trying to reach its goals.					
I'm unhappy with my team's level of commitment to the task.*					
The team members have conflicting aspirations for the class's performance.*					
This team does not give me enough opportunities to improve my personal performance.*					

*Items are reverse-scored.

Items measuring construction

	I disagree completely	I disagree somewhat	Neutral	I agree somewhat	I agree completely
In this team, I share all relevant information and ideas I have.					
Team members are listening carefully to each other.					

If something is unclear, we ask each other questions.					

Items measuring co-construction

	I disagree completely	I disagree somewhat	Neutral	I agree somewhat	I agree completely
Team members elaborate on each other's information and ideas.					
Information from team members is complemented with information from other team members.					
Team members draw conclusions from the ideas that are discussed in the team.					

Items measuring constructive conflict

	I disagree completely	I disagree somewhat	Neutral	I agree somewhat	I agree completely
My team tends to handle differences of opinions by addressing them directly.					
Comments on ideas are acted upon.					
Opinions and ideas of class members are verified by asking each other critical questions.					

Items measuring mutually shared cognition

	I disagree completely	I disagree somewhat	Neutral	I agree somewhat	I agree completely
At this moment, my team has a common understanding of the task we have to handle.					
At this moment, my team has a common understanding of how to deal with the task.					

Items measuring team effectiveness

	I disagree completely	I disagree somewhat	Neutral	I agree somewhat	I agree completely
I am satisfied with the performance of my team.					
We have completed the task in a way we all agree upon.					
I would want to work with my team in the future.					
As a team, we have learned a lot.					

Regarding the previous questions, is there anything you'd like to elaborate or clarify?

How did you select your partner for this course?

What barriers, if any, have you experienced in solving problems with your group?

After your group agrees on the correct answer, what do you typically do next?

What suggestions do you have for improving instruction or teamwork in this course?

What suggestions do you have for improving the lecture videos? Have you found them useful in your preparation? Why or why not?

Appendix D

Post-course survey (control and experimental students)

What are your full initials?

Who is your instructor? Please use the instructor's full name.

Following are six LSAT logical reasoning questions repeated from elsewhere in this course. Please answer each to the best of your ability, using the applicable methods and technique you have learned in this course. Feel free to use your own scratch if you choose.

1. Commentator: If a political administration is both economically successful and successful at protecting individual liberties, then it is an overall success. Even an administration that fails to care for the environment may succeed overall if it protects individual liberties. So far, the present administration has not cared for the environment but has successfully protected individual liberties.

   If all of the statements above are true, then which one of the following must be true?

   A. The present administration is economically successful.
   B. The present administration is not an overall success.
   C. If the present administration is economically successful, then it is an overall success.
   D. If the present administration had been economically successful, it would have cared for the environment.
   E. If the present administration succeeds at environmental protection, then it will be an overall success.

2. Baxe Interiors, one of the largest interior design companies in existence, currently has a near monopoly in the corporate market. Several small design companies have won prestigious awards for their corporate work, while Baxe has won none. Nonetheless, the corporate managers who solicit design proposals will only contract with companies they believe are unlikely to go bankrupt, and they believe that only very large companies are unlikely to go bankrupt.

   The statements above, if true, most strongly support which one of the following?

A. There are other very large design companies besides Baxe, but they produce designs that are inferior to Baxe's.
B. Baxe does not have a near monopoly in the market of any category of interior design other than corporate interiors.
C. For the most part, designs that are produced by small companies are superior to the designs produced by Baxe.
D. At least some of the corporate managers who solicit proposals are unaware that there are designs that are much better than those produced by Baxe.
E. The existence of interior designs that are superior to those produced by Baxe does not currently threaten its near monopoly in the corporate market.

3. A development company has proposed building an airport near the city of Dalton. If the majority of Dalton's residents favor the proposal, the airport will be built. However, it is unlikely that a majority of Dalton's residents would favor the proposal, for most of them believe that the airport would create noise problems. Thus, it is unlikely that the airport will be built.

   The reasoning in the argument is flawed in that the argument

   A. treats a sufficient condition for the airport's being built as a necessary condition
   B. concludes that something must be true, because most people believe it to be true
   C. concludes, on the basis that a certain event is unlikely to occur, that the event will not occur
   D. fails to consider whether people living near Dalton would favor building the airport
   E. overlooks the possibility that a new airport could benefit the local economy

4. Principle: When none of the fully qualified candidates for a new position art Arvue Corporation currently works for that company, it should hire the candidate who would be most productive in that position.

   Application: Arvue should not hire Krall for the new position, because Delacruz is a candidate and is fully qualified.

   Which one of the following, if true, justifies the above application of the principle?

   A. All of the candidates are fully qualified for the new position, but none already works for Arvue.
   B. Of all the candidates who do not already work for Arvue, Delacruz would be the most productive in the new potion.

C. Krall works for Arvue, but Delacruz is the candidate who would be most productive in the new position.
D. Several candidates currently work for Arvue, but Krall and Delacruz do not.
E. None of the candidates already works for Arvue, and Delacruz is the candidate who would be most productive in the new position.

5. There can be no individual freedom without the rule of law, for there is no individual freedom without social integrity, and pursuing the good life is not possible without social integrity.

The conclusion drawn above follows logically if which one of the following is assumed?

A. There can be no rule of law without social integrity.
B. There can be no social integrity without the rule of law.
C. One cannot pursue the good life without the rule of law.
D. Social integrity is possible only if individual freedom prevails.
E. There can be no rule of law without individual freedom.

6. Economists: Countries with an uneducated population are destined to be weak economically and politically, where those with an educated population have governments that display a serious financial commitment to public education. So any nation with a government that has mad e such a commitment will avoid economic and political weakness.

The pattern of flawed reasoning in which one of the following arguments is most similar to that in the economist's argument?

A. Animal species with a very narrow diet will have more difficulty surviving if the climate suddenly changes, but a species with a broader diet will not; for changes in the climate can remove the traditional food supply.
B. People incapable of empathy are not good candidates for public office, but those who do have the capacity for empathy are able to manipulate others easily; hence, people who can manipulate others are good candidates for public office.
C. People who cannot give orders are those who do not understand the personalities of the people to whom they give orders. Thus, those who can give orders are those who understand the personalities of the people to whom they give orders.

D. Poets who create poetry of high quality are those who have studied traditional poetry, because poets who have not studied traditional poetry are the poets most likely to create something shockingly inventive, and poetry that is shockingly inventive is rarely fine poetry.

E. People who dislike exercise are unlikely to lose weight without sharply curtailing their food intake; but since those who dislike activity generally tend to avoid it, people who like to eat but dislike exercise will probably fail to lose weight.

Please answer each of the following questions with regard to your experience during this course (questions displayed on following page).

	Strongly disagree	Somewhat disagree	Slightly disagree	Slightly agree	Somewhat agree	Strongly agree
"I make sure to study for this class on a regular basis."						
"I put forth effort in this class."						
"I do all of the homework problems for this class."						
"I look over class notes between classes to make I understand the material."						
"I take good notes in class."						
"I listen carefully in class."						
"I find ways to make the course material relevant to my life."						
"I apply course material to my life."						
"I find ways to make the course interesting to myself."						
"I think about the course between meetings."						
"I strongly desire to learn the course material."						
"I ask questions during class."						
"I enjoy attending class."						
"I help fellow students learn the course material."						

Works Cited

Astin, A. W. (1984). Student involvement: A developmental theory for higher education. *Journal of College Student Personnel, 25*, 297-308.

Astin, A. W. (1993). *What matters in college? Four critical years revisited.* San Fransisco: Jossey-Bass.

Baker, J. W. (2000). The "classroom flip": Using web course management tools to become the guide by the side. In J. A. Chambers (Ed.), *Selected Papers from the 11th International Conference on College Teaching and Learning* (pp. 9-17). Jacksonville: Florida Community College.

Barron, B. (2000). Acheiving coordination in collaborative problem-solving groups. *The Jounrnal of the Learning Sciences, 9*(4), 403-36.

Barron, B. (2003). When smart groups fail. *The Journal of the Leaning Sciences, 12*(3), 307-359.

Berkowitz, R. (1998). One point on the LSAT: How much is it worth? Standardized tests as a determinant of earnings. *The American Economist, 42*(2), 80-89.

Bishop, J. L., & Verleger, M. A. (2013). *The flipped classroom: A survey of the research.* 120th ASEE annual conference presentation, paper ID #6219. Atlanta: American Society for Engineering Education.

Bonwell, C. C., & Eison, J. A. (1991). *Active learning: Creating excitement in the classroom.* Washington, DC: School of Education and Human Devlopment, George Washington University.

Boud, D., Cohen, R., & Sampson, J. (1999). Peer learning and assessment. *Assessment & Evaluation in Higher Education, 24*(4), 413-426.

Bray, J. N. (2002). Uniting teacher learning: Collaborative inquiry for professional development. *New Directions for Adult and Continuing Education, 94*, 83-92. doi:10.1002/ace.62

Brophy, J. (1999). Toward a model of value aspects of motivation in education: Developing appreciation for particular learning domains and activities. *Educational Psychologist, 34*(2), 75-85.

Brown, A. F. (2012). *A phenomenological study of undergraduate instructors using the inverted or flipped classroom model.* Pepperdine University: Dissertation.

Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin*, 3-7.

Colbeck, C. L., Campbell, S. E., & Bjorklund, S. A. (2000, Jan.-Feb.). Grouping in the dark: What college students learn from group projects. *The Journal of Higher Education, 71*(1), 60-83.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences, 13*(1), 15-42.

Cross, T., & Slater, R. B. (1997). Special Report: Why the end of affirmative action would exclude all but a very few blacks from America's leading universities and graduate schools. *Journal of Blacks in Higher Education, 8*, 11-13.

Crouch, C. H., & Mazur, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics, 69*(9), 970-977.

Davies, R. S., Dean, D. L., & Ball, N. (2013). Flipping the classroom and instructional technology integration in a college-level information systems spreadsheet course. *Education Tech Research*, 563-580.

Didier, E. (2013, July 11). *Law School Predictor Full Time Programs*. Retrieved from Law School Predictor: http://www.lawschoolpredictor.com/wp-content/uploads/Law-School-Predictor-Full-Time-Programs.htm

Dirkx, J. M., & Jha, L. R. (1994). Completion and attrition in adult basic education. *Adult Education Quarterly*, 269-285.

Edmonds, C. T., & Edmonds, T. P. (2008). An empirical investigation of the effects of SRS technology on Introductory Managerial Accounting students. *Issues in Accounting Education, 23*(3), 412-434.

Evans, J., Sweeney, A. T., & Reese, L. M. (October 2011). *Summary of self-reported methods of test preparation by LSAT takers for testings years 2008-2009 through 2010-2011*. Newtown, PA: Law School Admission Council.

Feichtner, S. B., & Davis, E. A. (1984-85). Why some groups fail: A survey of students' experiences with learning groups. *Organizational Behavior Teaching Review, 9*(4), 58-71.

Fuchs, L. S., Fuchs, D., Kazdan, S., Karns, K., Calhoon, M. B., Hamlett, C. L., & Hewlett, S. (2000). Effects of workgroup structure and size on student productivity during collaborative work on complex tasks. *The Elementary School Journal, 100*(3), 182-212.

Giles, R. M., & Ashman, A. F. (Eds.). (2003). *Co-operative learning: The social and intellectual outcomes of learning in groups.* New York, NY: RoutledgeFalmer.

Guerrero, C. (2009). *Classroom response systems: Using clickers and collaboration to foster learning in a postsecondary setting.* University of Texas at Austin: Unpublished prospectus.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 1-26.

Haladyna, T., Haas, N., & Allison, J. (1998). Continuing tensions in standardized testing. *Childhood Education, 74*(5), 262-273. Retrieved from http://search.proquest.com.ezproxy.lib.utexas.edu/docview/1290009925?accountid=7118

Handelsman, M. M., Briggs, W. L., Sullivan, N., & Towler, A. (2005). A measure of college student course engagement. *The Journal of Educational Research, 98*(3), 184-191. Retrieved from http://www.jstor.org/stable/27548076

Harris, C. (2003). *Understanding the role of epistemological beliefs in post-graduate studies: Motivation and conceptions of learning in first-year law students.* University of Texas at Austin.

Johnson, D. W., & Johnson, R. T. (1998). Making cooperative learning work. *Theory into Practice, 38*(2), 67-73. Retrieved from http://www.jstor.org/stable/1477225?origin=JSTOR-pdf

Johnson, D. W., Johnson, R. T., & Smith, K. (1991). *Cooperative learning: Increasing college faculty instructional productivity (ASHE-ERIC Higher Education Rep. No. 91-1).* Washington, DC: Association for the Study of Higher Education.

Johnston, C. G., James, R. H., Lye, J. N., & McDonald, I. M. (2000). An evaluation of collaborative problem solving for learning economics. *The Journal of Economic Education, 31*(1), 13-29.

Klein, S. P., & Hamilton, L. (1998). *The validity of the U.S. News and World Report law school rankings.* Washington, D.C.: American Association of Law Schools.

Kuh, G. D., & Vesper, N. (1997). A comparision of student experiences with good practices in undergraduate education between 1990 and 1994. *Review of Higher Education, 21*, 43-61.

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education, 31*(1), 30-43.

Mackey, A. P., Whitaker, K. J., & Bunge, S. A. (2012). Experience-dependent plasticity in white matter microstructure: Reasoning training alters structural connectivity. *Frontiers in Neuroanatomy*, 1-9. doi:10.3389/fnana.2012.00032

Mazur, E. (1997). *Peer Instruction: A Users' Manual.* Upper Saddle River, NJ: Prentice Hall.

McConnell, D. (1999). Examining a collaborative assessment process in networked lifelong learning. *Journal of Computer Assisted Learning, 15*, 232-243.

McKeachie, W. J., & Svinicki, M. (2006). *McKeachie's Teaching Tips.* Boston: Houghton Mifflin Company.

McNeil, L. M. (2000). *Contradictions of school reform: Educational costs of standardized testing.* New York: Routledge.

Michaelsen, L. K., & Black, R. H. (1994). Building learning teams: The key to harnessing the power of small groups in higher education. In *Collaborative learning: A sourcebook for higher education* (Vol. 2, pp. 65-81).

Michaelsen, L. K., & Fink, L. D. (2008). Preface. In L. K. Michaelsen, M. Sweet, & D. X. Parmalee (Eds.), *Team-based learning: Small-group learning's next big step* (pp. i-vi). San Fransisco: Jossey-Bass.

Michaelsen, L. K., & Sweet, M. (2008). The essential elements of team-based learning. *New Directions for Teaching and Learning*(116), 7-27.

Michaelsen, L. K., & Sweet, M. (2011). Team-based learning. *New Directions for Teaching & Learning, 2011*(128), 41-51.

Michaelsen, L. K., Knight, A. B., & Fink, L. D. (Eds.). (2002). *Team-based learning: A transformative use of small groups.* Westport, CT: Praeger Publishers.

National Survey of Student Engagement. (2000). *The NSSE report: National benchmarks of effective educational practice.* Bloomington: Indiana University Center for Postsecondary Research and Planning.

Nystrand, M. (1997). *Opening dialogue: Understanding the dynamics of language and learning in the English Classroom.* New York, NY: Teachers College Press.

Oxford, R. L. (1997). Cooperative learning, collaborative learning, and interaction: Three communicative strands in the language classroom. *The Modern Language Journal, 81*(4), 443-456. Retrieved from http://www.jstor.org/stable/328888

Perin, D., & Greenberg, D. (1994, July). Understanding dropout in an urban worker education program. *Urban Education, 4*(3), 169-187.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology, 95*(4), 667-686.

Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications.* Upper Saddle River, NJ: Merrill Prentice-Hall.

Pipes, R. B., & Wilson, J. M. (1996). A multimedia model for undergraduate education. *Technology in Society, 18*(3), 387-401.

Randall, V. R. (1995-1996). The Myers-Briggs Type Indicator, first year law students and performance. *Cumberland Law Review*, 63-105.

Randall, V. R. (2006). The misuse of the LSAT: Discrimination against blacks and other minorities in law school admissions. *St. John's Law Review*.

Ravitch, D. (2010). *The death and life of the great American school system: How testing and choice are undermining education.* New York, NY: Basic Books.

Reinking, D., & Bradley, B. A. (2008). *Formative and Design Experiments: Approaches to Language and Literacy Research.* New York, NY: Teachers College Press.

Roschelle, J., & Teasley, S. D. (1995). Construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer-supported collaborative learning* (pp. 69-97). New York: Springer-Verlag.

Saleh, M., Lazonder, A. W., & de Jong, T. (2007). Structuring collaboration in mixed-ability groups to promote verbal interaction, learning, and motivation of average-ability students. *Contemporary Educational Psychology, 32*(3), 314-31.

Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. *The Journal of the Learning Sciences, 4*, 321-354.

Shirouzu, H., Mikaye, N., & Masukawa, H. (2002). Cognitively active externalization for situated reflection. *Cognitive Science, 26*, 469-501.

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional paricipation in academic activities in the classroom. *Educational and Psychological Measurement, 69*(3), 493-525.

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

Slavin, R. E. (1996). Research for the future: Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21*, 43-69.

Smith, M. L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher,* 8-11.

Springer, L., Stanne, E. S., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematices, engineering and technology: A meta-analysis. *Review of Educational Research, 69*(21), 21-51.

Strayer, J. F. (2007). *The effects of the classroom flip on the learning envrionment: A comparison of learning activity in a traditional classroom and a flip classroom that used an intelligent tutoring system.* The Ohio State University: Dissertation.

Strayer, J. F. (2012). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learning Environment Research, 15*, 171-193.

Stuart, S., & Vance, R. (2013). Brining a knife to a gunfight: The academically underprepared law student & legal education reform. *Valparaiso Law Faculty Publications*. Retrieved from http://scholar.valpo.edu/law_fac_pubs/116/.

Svanum, S., & Bigatti, S. M. (2009). Academic course engagement during one semester forecasts college success: Engaged students are more likely to earn a degree, do it faster, and do it better. *Journal of College Student Development,* 120-132.

Van den Bossche, P., Gijsalears, W. H., Segers, M., & Kirschner, P. (2006, October). Social and cognitive factors driving teamwork in collaborative learning environments: Team learning beliefs and behaviors. *Small Group Research, 37*(5), 490-521.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, Massachusetts: Harvard University Press.

Webb, N. M., Nemer, K. M., Chizhik, A. W., & Sugrue, B. (1998). Equity issues in collaborative group assessment: Group composition and performance. *American Educational Research Journal, 35*(4), 607-51.

Woolf, B. P. (2010). *A roadmap for education technology.* Global Resources for Online Education.